

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2003, 17:26:01 ; Search time 32 Seconds
(without alignments)
231.803 Million cell updates/sec

Title: US-09-866-066-20

Perfect score: 201

Sequence: 1 PYKWPFFHFAFFHOILATSTIAMSHCRPTVYKQTSII 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_21.*

2: sp_archaea.*

3: sp_bacteria.*

4: sp_fungi.*

5: sp_human.*

6: sp_invertebrate.*

7: sp_mammal.*

8: sp_mhc.*

9: sp_organelle.*

10: sp_phase.*

11: sp_plant.*

12: sp_rodent.*

13: sp_virus.*

14: sp_vertebrate.*

15: sp_unclassified.*

16: sp_rviro.*

17: sp_bacteriap.*

18: sp_archaea.*

19: sp_bacteria.*

20: sp_fungi.*

21: sp_human.*

22: sp_invertebrate.*

23: sp_mammal.*

24: sp_mhc.*

25: sp_organelle.*

26: sp_phase.*

27: sp_plant.*

28: sp_rodent.*

29: sp_virus.*

30: sp_vertebrate.*

31: sp_unclassified.*

32: sp_rviro.*

33: sp_bacteriap.*

34: sp_archaea.*

35: sp_bacteria.*

36: sp_fungi.*

37: sp_human.*

38: sp_invertebrate.*

39: sp_mammal.*

40: sp_mhc.*

41: sp_organelle.*

42: sp_phase.*

43: sp_plant.*

44: sp_rodent.*

45: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 56.5 | 28.1 | 176 | 6 | 077545 sus scrofa |
| 2 | 55 | 27.4 | 82 | 4 | Q961S2 homo sapien |
| 3 | 55 | 27.4 | 802 | 10 | Q8RXV4 arabidopsis |
| 4 | 55 | 27.4 | 837 | 10 | Q80590 arabidopsis |
| 5 | 54 | 26.9 | 1031 | 10 | Q9C959 arabidopsis |
| 6 | 53.5 | 26.6 | 319 | 11 | Q8VGS1 mus musculus |
| 7 | 53 | 26.4 | 324 | 16 | Q8YSL1 anabaena sp |
| 8 | 53 | 26.4 | 343 | 10 | Q43100 caenorhabdi |
| 9 | 53 | 26.4 | 485 | 5 | Q18544 caenorhabdi |
| 10 | 53 | 26.4 | 560 | 10 | Q9LQ07 arabidopsis |
| 11 | 53 | 26.4 | 866 | 16 | Q8UCG6 agrobacteri |
| 12 | 52.5 | 26.1 | 223 | 11 | Q921U7 mus musculu |
| 13 | 52.5 | 26.1 | 319 | 11 | Q8VFL0 mus musculu |
| 14 | 52 | 25.9 | 343 | 10 | Q40949 populun nig |
| 15 | 52 | 25.9 | 427 | 16 | Q69687 mycobacteri |
| 16 | 51.5 | 25.6 | 355 | 10 | Q9AUL4 oryza sativ |

| | | | | | |
|----|------|------|------|----|--------|
| 17 | 51 | 25.4 | 227 | 8 | Q9B6X7 |
| 18 | 51 | 25.4 | 343 | 10 | Q40950 |
| 19 | 51 | 25.4 | 343 | 10 | Q43099 |
| 20 | 51 | 25.4 | 355 | 6 | Q8WRR5 |
| 21 | 51 | 25.4 | 355 | 6 | Q8WRR4 |
| 22 | 51 | 25.4 | 364 | 5 | Q76790 |
| 23 | 51 | 25.4 | 606 | 3 | Q88299 |
| 24 | 50.5 | 25.1 | 400 | 16 | Q8Y295 |
| 25 | 50 | 24.9 | 329 | 5 | Q17185 |
| 26 | 50 | 24.9 | 512 | 16 | Q8Y3R4 |
| 27 | 50 | 24.9 | 1150 | 10 | Q8YF70 |
| 28 | 49.5 | 24.6 | 188 | 12 | Q89518 |
| 29 | 49.5 | 24.6 | 263 | 5 | Q9YXR2 |
| 30 | 49 | 24.4 | 217 | 10 | Q9LXW4 |
| 31 | 49 | 24.4 | 320 | 4 | Q8WZ55 |
| 32 | 49 | 24.4 | 2849 | 4 | Q8TDX9 |
| 33 | 48.5 | 24.1 | 611 | 8 | Q9T6Z0 |
| 34 | 48 | 23.9 | 137 | 16 | Q93409 |
| 35 | 48 | 23.9 | 247 | 10 | Q9LN66 |
| 36 | 48 | 23.9 | 270 | 16 | Q9KT16 |
| 37 | 48 | 23.9 | 346 | 10 | Q93VA1 |
| 38 | 48 | 23.9 | 425 | 16 | Q986U6 |
| 39 | 48 | 23.9 | 522 | 16 | Q9CLG3 |
| 40 | 48 | 23.9 | 524 | 5 | Q21729 |
| 41 | 48 | 23.9 | 567 | 5 | Q8T3D4 |
| 42 | 48 | 23.9 | 583 | 5 | Q96395 |
| 43 | 48 | 23.9 | 640 | 16 | Q915U9 |
| 44 | 48 | 23.9 | 729 | 5 | Q9W0J0 |
| 45 | 48 | 23.9 | 1455 | 10 | Q9AXC3 |

ALIGNMENTS

RESULT 1

077545 PRELIMINARY; PRT; 176 AA.
AC 077545;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Olfactory receptor (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98346353; PubMed=9682823;
RA Matarazzo V., Tirard A., Renucci M., Belaich A., Clement J.L.;
RT *Isolation of putative olfactory receptor sequences from pig nasal epithelium.*
RL Neurosci. Lett. 249:87-90(1998).
DR EMBL; AF042019; AAC24795.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEPTOR_F1_2; 1.
KW Receptor.
FT NON_TER 1
FT NON_TER 176
SQ SEQUENCE 176 AA; 19664 MW; 3678231F7F0DD3F8 CRC64;

Query Match 28.1%; Score 56.5; DB 6; Length 176;
Best Local Similarity 37.9%; Pred. No. 1.2;
Matches 11; Conservative 4; Mismatches 11; Indels 3; Gaps 1;

Oy 1 PYKWPFFHFAFFHOILATSTIAMSHCRPTV 29

Db 33 PYSYFSMALFHTIL---TFRLSYCHSNI 58

RESULT 2

Q961S2

DR Pfam: PF00271; hellicase_C; 1.
 DR Pfam: PF00570; HRDC; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELICC; 1.
 DR SMART: SM00341; HRDC; 1.
 DR TIGRPFAMs: TIGR00614; recq; 1.
 DR PROSITE: PS00690; DEAH_LTP_HELICASE; UNKNOWN_1.
 DR PROSITE: PS00453; FKBP_LTP_PHASE_1; UNKNOWN_1.
 KW ATP-binding; Hellicase.
 SQ SEQUENCE 1031 AA; 115790 MW; A57D7A39330F251F CRC64;

Query Match 26.9%; Score 54; DB 10; Length 1031;
 Best Local Similarity 46.2%; Pred. No. 17;
 Matches 12; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 10 FPHQILATSTAMSHCRPTVYKQTSI 35
 DB 170 FQAIIIDVDQIAMEHCPCSTCTKQPSV 195

RESULT 6

Q8VGS1 PRELIMINARY; PRT; 319 AA.
 AC Q8VGS1;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Olfactory receptor MOR185-3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang X., Firestein S.J.;
 RT "The olfactory receptor gene superfamily of the mouse."
 RL Nat. Neurosci. 0:0-0(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Adams M.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY073075; AAL60738.1; -.
 DR InterPro: IPR000276; GPCR_Rhodospn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 319 AA; 36135 MW; 9452501AFBB25172 CRC64;

Query Match 26.6%; Score 53.5; DB 11; Length 319;
 Best Local Similarity 37.9%; Pred. No. 6.1;
 Matches 11; Conservative 5; Mismatches 10; Indels 3; Gaps 1;

QY 1 PYKPFHFAFFHQLATSTAMSHCRPTV 29
 DB 148 PYVSELVAFHTII---TFELTYCGPNV 173

RESULT 7

Q8VSL1 PRELIMINARY; PRT; 324 AA.
 AC Q8VSL1;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein Alr3073.
 GN ALR3073.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21595285; PubMed=11759840;

RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL: AP003591; BAB74772.1; -.
 DR InterPro: IPR001173; Glycos_transf_2.
 DR Pfam: PF00535; Glycos_transf_2; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 324 AA; 37168 MW; 0BDA5EF1013562D6 CRC64;

Query Match 26.4%; Score 53; DB 16; Length 324;
 Best Local Similarity 29.4%; Pred. No. 7.4;
 Matches 10; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

QY 2 YKPFHFAFFHQLATSTAMSHCRPTVYKQTSI 35
 DB 185 YKPFVVVKHQILYRVSVTSKFQLIIRASL 218

RESULT 8

Q43100 PRELIMINARY; PRT; 343 AA.
 AC Q43100;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Peroxidase (EC 1.11.1.7).
 OS Populus trichocarpa (Western balsam poplar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids 1; Malpighiales; Salicaceae; Populus.
 OX NCBI_TaxID=3694;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-VXLEM;
 RA Christensen J.H., Bauw G., Boerjan W., Van Montagu M.;
 RT "Purification and cloning of poplar xylem peroxidases.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X97349; CAA66035.1; -.
 DR HSSP: P00433; 2ATJ.
 DR InterPro: IPR002016; Peroxidase.
 DR Pfam: PF00141; peroxidase; 1.
 DR PRINTS: PR00458; PEROXIDASE.
 DR PROSITE: PS00435; PEROXIDASE_1; 1.
 DR PROSITE: PS00436; PEROXIDASE_2; 1.
 KW Oxidoreductase; Peroxidase.
 SQ SEQUENCE 343 AA; 36568 MW; C55A6BE43349D50B CRC64;

Query Match 26.4%; Score 53; DB 10; Length 343;
 Best Local Similarity 44.0%; Pred. No. 7.8;
 Matches 11; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 9 AFFHQILATSTAMSHCRPTVYKOT 33
 DB 10 AFFVLLRGTLACGQLTPFYDQT 34

RESULT 9

Q18544 PRELIMINARY; PRT; 485 AA.
 AC Q18544;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE C39E9.7 protein.
 GN C39E9.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.

DR PROSITE; PS00237; G_PROTEIN_RECEPT_F1_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEPT_F1_2; 1.
 KW RECEPTOR.
 FT NON_TER 1
 FT NON_TER 223
 SQ SEQUENCE 223 AA; 25018 MW; B6AE52CEB8863A34 CRC64;
 Query Match 26.1%; Score 52.5; DB 11; Length 223;
 Best Local Similarity 25.0%; Pred. No. 6;
 Matches 14; Conservative 6; Mismatches 13; Indels 23; Gaps 2;
 QY 1 PYKWPFFHFAFHQILATSTIAMSHCRPTV-----YKQTSII 36
 DB 85 PYSYSLMALIHILL---TFRLSYCHSNIINHFYCDMPLLRLTCSATHYKQLSIL 137
 RESULT 13
 ID Q8VFL0 PRELIMINARY; PRT; 319 AA.
 AC Q8VFL0;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DE Olfactory receptor MOR185-6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang X., Firestein S.J.;
 RT "The olfactory receptor gene superfamily of the mouse.";
 RL Nat. Neurosci. 0:0-0(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Adams M.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY073514; AAL61177.1;
 DR InterPro; IPR000276; GPCR_Rhodopsin.
 DR Pfam; PF00001; 7tm1.1;
 DR PRINTS; PR00237; GPCRHHODOPS.
 DR PROSITE; PS00237; G_PROTEIN_RECEPT_F1_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEPT_F1_2; 1.
 KW RECEPTOR.
 SQ SEQUENCE 319 AA; 36289 MW; 4D1261ABD0B15033 CRC64;
 Query Match 26.1%; Score 52.5; DB 11; Length 319;
 Best Local Similarity 25.0%; Pred. No. 8.6;
 Matches 14; Conservative 6; Mismatches 13; Indels 23; Gaps 2;
 QY 1 PYKWPFFHFAFHQILATSTIAMSHCRPTV-----YKQTSII 36
 DB 148 PYSYSLMALIHILL---TFRLSYCHSNIINHFYCDMPLLRLTCSATHYKQLSIL 200
 RESULT 14
 ID Q40949 PRELIMINARY; PRT; 343 AA.
 AC Q40949;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DE Peroxidase.
 OS Populus nigra (Lombardy poplar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Salicaceae; Populus.
 OX NCBI_TaxID=3691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kato Y., Sakuma Y., Azuma T., Ando A., Miura K., Takabe K., Tamai Y.;
 RT "Isolation and characterization of a novel anionic peroxidase cDNA
 found in poplar (Populus nigra) suspension cultured cells.";

RL J. Wood Sci. 47:165-170(2001).
 DR EMBL; D83224; BAA11852.1;
 DR HSSP; P00433; 2ATJ.
 DR InterPro; IPR002016; Peroxidase.
 DR Pfam; PF00141; peroxidase.1.
 DR PRINTS; PR00458; PEROXIDASE.
 DR PROSITE; PS00435; PEROXIDASE_1; 1.
 DR PROSITE; PS00435; PEROXIDASE_2; 1.
 SQ SEQUENCE 343 AA; 36529 MW; B4A540ADDFAF91AA CRC64;
 Query Match 25.9%; Score 52; DB 10; Length 343;
 Best Local Similarity 44.0%; Pred. No. 11;
 Matches 11; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
 QY 9 AFFHQILATSTIAMSHCRPTVYKOT 33
 DB 10 AFFEVLGGTLAGQLTPTFYDOT 34
 RESULT 15
 ID O69687 PRELIMINARY; PRT; 427 AA.
 AC O69687;
 DT 01-AUG-1998 (TRENBLrel. 07, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Hypothetical 46.9 kDa protein (Cyclopropane-fatty-acyl-phospholipid
 synthase).
 DE GN RV3720 OR MT3823 OR MTV025.068.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OSHKOSH;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Boucher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL022121; CAAL8042.1; ALT_INIT.
 DR EMBL; AE007178; AAK48192.1;
 DR TIGR; MT3823;
 DR Tuberculist; RV3720;
 DR InterPro; IPR003333; CMAS.
 DR InterPro; IPR000051; SAM_bind.
 DR Pfam; PF02353; CMAS; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 427 AA; 47600 MW; 5830E8D70E4822C5 CRC64;
 Query Match 25.9%; Score 52; DB 16; Length 427;
 Best Local Similarity 36.4%; Pred. No. 14;
 Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Wed Mar 26 09:14:09 2003

us-09-866-066-20.rspt

Page 6

QY 4 WDPHAFPHQILATSTIANSHC 25
Db :||:| | | : :||
281 YPFYEGFLKSLRTGGLLNHC 302

Search completed: March 24, 2003, 17:26:53
Job time : 35 secs

OX NCBI_TaxID=6239;
 RN SEQUENCE FROM N.A.
 RP Sims M.A.;
 RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RX MEDLINE=99069613; PubMed=9851916;
 RP SEQUENCE FROM N.A.
 RA none;
 RX MEDLINE=21608550; PubMed=11743193;
 RP "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; 270307; CRA94336.1;
 SQ SEQUENCE 485 AA; 55578 MW; 4AA2B45A957AFC29 CRC64;

Query Match 26.4%; Score 53; DB 5; Length 485;
 Best Local Similarity 40.7%; Pred. No. 11;
 Matches 11; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 7 HPAFFHQILATSTIAMSHCRPTVYKQT 33
 DB 60 HIPFFEFILCTLCMASITSNPVIQST 86

RESULT 10
 ID Q9LQO7 PRELIMINARY; PRT; 560 AA.
 AC Q9LQO7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 GN F2489.5
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN SEQUENCE FROM N.A.
 RP STRAIN=CV. COLUMBIA;
 RA Liu S., Yu G., Sakano H., Jhaveri A., Lee J., Lenz C., Pham P.,
 RA Toriumi M., Chin C., Choi E., Chlou J., Gonzalez A., Chung M.,
 RA Hong B., Koo T., Li J., Liu A., Vaysberg M., Altafi H., Brooks S.,
 RA Buchler E., Chao Q., Conn L., Conway A.B., Hansen N.,
 RA Johnson-Hopson C., Khan S., Kim C., Lam B., Nguyen M., Palm C.,
 RA Shlan P., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
 RT "The sequence of BAC F2489 from Arabidopsis thaliana chromosome 1."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007583; AAF75069.1;
 SQ SEQUENCE 560 AA; 63994 MW; A233433AA854FA77 CRC64;

Query Match 26.4%; Score 53; DB 10; Length 560;
 Best Local Similarity 33.3%; Pred. No. 13;
 Matches 14; Conservative 5; Mismatches 17; Indels 6; Gaps 2;

QY 1 PKWPFHFAFHQILATSTI-----AMSHCRP--TVYKQTSII 36
 DB 47 PLPMLFIFLFLPLLLISTGGGSRKTRCPSSSTYSLSLV 88

RESULT 11
 ID Q8UCG6 PRELIMINARY; PRT; 866 AA.
 AC Q8UCG6;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE GN Hypothetical protein Atu2521.
 OS ATU2521 OR AGR_C.4580.
 OC Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OX NCBI_TaxID=176299;
 RN SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deatherage C., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Middle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58."
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Ourliio B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58."
 RL Science 294:2323-2328(2001).
 DR EMBL; AE009198; AAL43508.1;
 DR EMBL; AE008164; AAK8252.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 866 AA; 93876 MW; 1D21156AA010EFCE CRC64;

Query Match 26.4%; Score 53; DB 16; Length 866;
 Best Local Similarity 38.1%; Pred. No. 20;
 Matches 16; Conservative 4; Mismatches 12; Indels 10; Gaps 2;

QY 4 WPFHF-----AFFHQILATSTIAM-SHCRPTVYKQTSI 35
 DB 496 WEFESEAPRGLRALLGLVLASSTIAFSLMRPVTFPDSI 537

RESULT 12
 ID Q92107 PRELIMINARY; PRT; 243 AA.
 AC Q92107;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Olfactory receptor E6 (Fragment).
 GN OLF52.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BL/6; TISSUE=OLFACORY EPITHELIUM;
 RX MEDLINE=99091050; PubMed=9875846;
 RA Krautwurst D., Iau K.W., Reed R.N.;
 RT "Identification of ligands for olfactory receptors by functional
 expression of a receptor library."
 RL Cell 95:917-926(1998).
 DR EMBL; AF102528; AAD13320.1;
 DR MGD; MGI:1333748; Olf52.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.

T2711.1.
Arabidopsis thaliana (Mouse-ear cress).
GN Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OS Spermatophyta; Magnoliophyta; eudicotyledons; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
NCBI_TaxID=3702;
RN [1]
RR SEQUENCE FROM N.A.
RP STRAIN=CV. COLOMBIA;
RC Federspiel N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,
RA Au M., Araujo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,
RR Oji O., Osborne B.I., Shinn P., Sun H., Toriumi M., Vysotskaia V.S.,
RA Yu G., Ecker J., Theologis A., Davis R.W.;
RRL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
BL EMBL; AC004122; AAC34330.1; --
DR InterPro: IPR000504; RNA_rec_mot.
DR PROSITE: PS00030; RRM_RNP_1; UNKNOWN_1.
SQ SEQUENCE 837 AA; 95999 MW; A569462111C1AB5D CRC64;

Query Match 27.4%; Score 55; DB 10; Length 837;
Best Local Similarity 31.2%; Pred. No. 9.6;
Matches 10; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

OY 5 PFHFAFFHQILATSTAMSHCRPTVYKQTSII 36
DB 190 PVHFDTHFHAVLDVSVRHSVMKSAAYKRPAI 221
||||| : : : : : || : : :
||| ||| : : : : : || : :

RESULT 5
QC959 PRELIMINARY; PRT; 1031 AA.
ID QC9359;
AC QC9359;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative DNA helicase.
OS TPI.1.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eucotids; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
NCBI_TaxID=3702;
RN [1]
RR SEQUENCE FROM N.A.
RP STRAIN=CV. COLOMBIA;
RC MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetskata I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li X.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziali A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizso M., Rooney T., Rowley D.,
RA Sakano H., Salterberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana";
RL Nature 408:816-820(2000).
RL EMBL; AC018908; AAGS1646.1; --
DR InterPro: IPR001410; DEAH.
DR InterPro: IPR002464; DEAH_box.
DR InterPro: IPR001179; FKBP_Pptase.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR002121; HRDC.
DR InterPro: IPR004589; RecQ.
DR Pfam; PF00270; DEAD; 1.

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2003, 17:26:21 ; Search time 17 Seconds
(without alignments)
203.579 Million cell updates/sec

Title: US-09-866-066-20

Perfect score: 201

Sequence: 1 PYKWPFFHFAFFHQILATSTIAMSHCRPTVYKQTSII 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 67 | 33.3 | 1001 | 2 T13807 | potassium channel |
| 2 | 55 | 27.4 | 837 | 2 T00618 | hypothetical prote |
| 3 | 54 | 26.9 | 1031 | 2 G96634 | probable DNA helic |
| 4 | 53 | 26.4 | 324 | 2 AB2190 | hypothetical prote |
| 5 | 53 | 26.4 | 485 | 2 T19853 | hypothetical prote |
| 6 | 53 | 26.4 | 560 | 2 A86214 | hypothetical prote |
| 7 | 53 | 26.4 | 866 | 2 C97662 | hypothetical prote |
| 8 | 53 | 26.4 | 866 | 2 AF2886 | conserved hypothet |
| 9 | 52 | 25.9 | 343 | 2 T09565 | peroxidase (EC 1.1 |
| 10 | 52 | 25.9 | 420 | 2 A70796 | probable cycloprop |
| 11 | 51 | 25.4 | 334 | 2 T19860 | hypothetical prote |
| 12 | 51 | 25.4 | 343 | 2 T09566 | peroxidase (EC 1.1 |
| 13 | 51 | 25.4 | 364 | 2 T43361 | probable potassium |
| 14 | 51 | 25.4 | 606 | 2 S66857 | probable membrane |
| 15 | 50.5 | 25.1 | 400 | 2 AG1879 | hypothetical prote |
| 16 | 50 | 24.9 | 329 | 2 T43509 | probable potassium |
| 17 | 50 | 24.9 | 336 | 2 T32347 | outward rectifier |
| 18 | 50 | 24.9 | 512 | 2 AD1132 | transcription regu |
| 19 | 49.5 | 24.6 | 213 | 2 D69938 | hemolysin III homo |
| 20 | 49.5 | 24.6 | 263 | 2 F87756 | protein T01A4.2 [1 |
| 21 | 49 | 24.4 | 217 | 2 T48949 | hypothetical prote |
| 22 | 49 | 24.4 | 555 | 2 S63177 | mannosyl transfera |
| 23 | 49 | 24.4 | 1083 | 2 H86921 | probable arabinosy |
| 24 | 48.5 | 24.1 | 485 | 2 B65072 | probable transport |
| 25 | 48.5 | 24.1 | 485 | 2 C91098 | probable transport |
| 26 | 48.5 | 24.1 | 485 | 2 G85943 | probable transport |
| 27 | 48 | 23.9 | 220 | 2 F34012 | H+-transporting tw |
| 28 | 48 | 23.9 | 220 | 2 E34012 | H+-transporting tw |
| 29 | 48 | 23.9 | 220 | 2 D34012 | H+-transporting tw |

H+-transporting tw
H+-transporting tw
H+-transporting tw
periplasmic bindin
neuropeptide Y rec
hypothetical prote
conserved hypothet
homogenisate 1,2-
maturase-related p
probable arabinosy
hypothetical prote
hypothetical prote
protein C50H11.11
probable G protein
genome polyprotein
genome polyprotein

ALIGNMENTS

RESULT 1

T13807
potassium channel protein - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: T13807
R:Goldstein, S.A.; Price, L.A.; Rosenthal, D.N.; Pausch, M.H.
Proc. Natl. Acad. Sci. U.S.A. 93, 13256-13261, 1996
A:Title: ORK1, a potassium-selective leak channel with two pore domains cloned from *D*
A:Reference number: Z17770; MUID:97075152; PMID:8917578
A:Accession: T13807
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1001 <GOL>
A:Cross-references: EMBL:U55321; NID:g3808067; PID:g3808068; PIDN:AAC69250.1
C:Genetics:
A:Gene: ORK1
A:Cross-references: FlyBase:Fbgn0017561
A:Map position: 1

Query Match

Best Local Similarity 33.3%; Score 67; DB 2; Length 1001;

Matches 11; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY 1 PYKWPFFHFAFFHQILATSTIAMSHCRPTVY 30
||| | : ||| : ||| : ||| :
Db 87 PYTWTFYHAFFAFTVCVGYGNISPTTF 116

RESULT 2

T00618
hypothetical protein T2711.1 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 24-Nov-1999

C:Accession: T00618

R:Federspiel, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Arau

submitted to the EMBL Data Library, September 1998

A:Reference number: Z14193

A:Accession: T00618

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-837 <FED>

A:Cross-references: EMBL:AC004122; NID:g3176693; PID:g3540180; GSPOB:GN00059; ATSP:T2

A:Experimental source: cultivar Columbia

C:Genetics:

A:Gene: ATSP:T2711.1

A:Map position: 1

A:Introns: 58/2; 87/1; 138/3; 223/2; 241/3; 321/3; 366/2; 427/3; 540/3; 627/1;

C:Superfamily: *Arabidopsis thaliana* hypothetical protein T2711.1

Query Match

27.4%; Score 55; DB 2; Length 837;

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*
 A:Reference number: A97359; PMID:11743194
 A:Accession: C97662
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-866 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK88252.1; PID:g15157710; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C_4580
 A:Map position: circular chromosome

Query Match 26.4%; Score 53; DB 2; Length 866;
 Best Local Similarity 38.1%; Pred. No. 16;
 Matches 16; Conservative 4; Mismatches 12; Indels 10; Gaps 2;

OY 4 WPFHF-----AFFHQILATSTIAM-SHCRPTVYKQTSI 35
 DB 496 WEFEFSEAPRGLRALLGLVLAASSTIAIFSLMRPVTFKPD SI 537

RESULT 8
 AF2886
 conserved hypothetical protein Atu2521 [Imported] - *Agrobacterium tumefaciens* (strain C58)
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C:Accession: AF2886
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, K.; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AF2886
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-866 <KUR>
 A:Cross-references: GB:AE008688; PIDN:AAL43508.1; PID:g17741016; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu2521
 A:Map position: circular chromosome

Query Match 26.4%; Score 53; DB 2; Length 866;
 Best Local Similarity 38.1%; Pred. No. 16;
 Matches 16; Conservative 4; Mismatches 12; Indels 10; Gaps 2;

OY 4 WPFHF-----AFFHQILATSTIAM-SHCRPTVYKQTSI 35
 DB 496 WEFEFSEAPRGLRALLGLVLAASSTIAIFSLMRPVTFKPD SI 537

RESULT 9
 T09565
 peroxidase (EC 1.11.1.7) - black poplar
 C:Species: *Populus nigra* (black poplar)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
 C:Accession: T09565
 R:Sakuma, Y.; Azuma, T.; Kato, Y.; Kojima, Y.; Miura, K.
 submitted to the EMBL Data Library, January 1996
 A:Description: Poplar peroxidase gene responding to cytokinin treatment.
 A:Reference number: Z16737
 A:Accession: T09565
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-343 <SAK>
 A:Cross-references: EMBL:D83224
 A:Experimental source: strain italica Koehne
 C:Superfamily: peroxidase
 C:Keywords: chromoprotein; heme; iron; oxidoreductase

Query Match 25.9%; Score 52; DB 2; Length 343;

Best Local Similarity 44.0%; Pred. No. 8.3;
 Matches 11; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

OY 9 AFFHQILATSTIAMSHCRPTVYKOT 33
 DB 10 AFFFVLLGGTILACGQLTPTFYDQT 34

RESULT 10
 A70796
 probable cyclopropane-fatty-acyl-phospholipid synthase (EC 2.1.1.79) RV3720 [similar]
 C:Species: *Mycobacterium tuberculosis*
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 21-Jul-2000
 C:Accession: A70796
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Garnier, C.; Harris, D.; Gordon-Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: A70796
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-420 <COL>
 A:Cross-references: GB:AL022121; GB:AL123456; NID:g3261559; PIDN:CAA18042.1; PID:g296
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV3720
 C:Superfamily: cyclopropane-fatty-acyl-phospholipid synthase; MCM homology
 C:Keywords: methyltransferase; phospholipid metabolism; S-adenosylmethionine

Query Match 25.9%; Score 52; DB 2; Length 420;
 Best Local Similarity 36.4%; Pred. No. 10;
 Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 4 WPFHF-----AFFHQILATSTIAMSHC 25
 DB 274 YPFYFGLKSLTGGLLNHC 295

RESULT 11
 T19860
 hypothetical protein C40C9.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T19860
 R:Hembry, C.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z19188
 A:Accession: T19860
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-334 <WIL>
 A:Cross-references: EMBL:Z70266; PIDN:CAA94204.1; GSPDB:GN00028; CESP:C40C9.1
 A:Experimental source: clone C40C9
 C:Genetics:
 A:Gene: CESP:C40C9.1
 A:Map position: X
 A:Introns: 34/1; 60/2; 98/1; 145/3; 160/3; 181/1; 204/1; 252/2; 279/2; 306/3

Query Match 25.4%; Score 51; DB 2; Length 334;
 Best Local Similarity 29.6%; Pred. No. 11;
 Matches 8; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

OY 2 YKWPFPFAFFHQILATSTIAMSHCRPT 28
 DB 79 YQWQFAGAFATVTVITVGYGHSAPS 105

RESULT 12
 T09566
 peroxidase (EC 1.11.1.7) - black poplar

| | Best Local Similarity | 31.2% | Pred. No. 7.7; | | |
|----|-----------------------|-----------------------------------|----------------|----|----------------------------------|
| | Matches | 10; | Conservative | 8; | Mismatches 14; Indels 0; Gaps 0; |
| QY | 5 | PFHFAPFHQILATSTAMSHCRPTVYKQTSII | 36 | | |
| | | | : | : | : |
| Db | 190 | PVHFEDTFHAYLIDVSVHVSVMKAAAYKRPAIL | 221 | | |
| | | | : | : | : |

RESULT 3
G96634
probable DNA helicase w7P1.7 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G96634
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ensen, N.F.; Hughes, B.; Huizlar, L.

Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Sako, H.; Saito, R.; Marsiali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G96634
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1031 <STO>
A:Cross-references: GB:AE005173; NID:G6751700; PIDN:AAF27683.1; GSPOB:GNO0141
C:Genetics:
A:Gene: T7p1.7
A:Map position: 1

```

Query Match      25.98; Score 54; DB 2; Length 1031;
Best Local Similarity 46.2; Pred. No. 13;
Matches 12; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

Qy 10 FFHQILATSIAMSHCRPTVYKQTSI 35
      ||| ||| ||| |||
Db 170 FGAQIIDVDIAMEHCPSCTCKQPSV 195

```

RESULT 4
AB2190
hypothetical protein alr3073 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AB2190
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807. MUID:21595285. PMID:11759840

A:Residues: 1-324 <KUR>
A:Cross-references: GB:BA000019: PIDN:BAB74772.1: PID:g17132167: GSPDB:GN00179
A:Experimental source: strain FCC 7120
C:Genetics:
A:Gene: alr3073

| | | | | |
|-----------------------|--------------|----------------|----------------|-------------|
| Query Match | 26.4% | Score 53; | DB 2; | Length 324; |
| Best Local Similarity | 29.4%; | Pred. No. 5.6; | | |
| Matches 10; | Conservative | 7; | Mismatches 17; | Indels 0; |
| | | | | Gaps 0; |

QY 2 YKWPPEFAFFHQILATSIAMSHCRPTVYKQTSI 35
 ||||| |||| : : : :
 DB 185 YKWPFWVVKQHOILYRRS VTSKSFKLOIREASL 218

185 YKWPFWVKOHQOILYRRSVTSKSFKLOIREASL 218

RESULT 5
T19853
hypothetical protein C39E9.7 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T19853
R;Sims, M.
submitted to the ENBL Data Library, March 1996

A;Accession: preliminary; translated from GB/EMBL/DDBJ
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-485 <MIL>
A;Cross-references: EMBL:Z70307; PIN:CAA94336.1; GSPDB:GN00022; CESP:C39E9.7
A;Experimental source: clone C39E9
C;Genetics:
A;Gene: CESP:C39E9.7
A;Map position: 4
A;Introns: 30/1; 59/3; 101/1; 160/2; 197/1; 288/3; 370/3; 431/1

```

Query Match      26.48; Score 53; DB 2; Length 485;
Best Local Similarity 40.7%; Pred. No. 8.5;
Matches 11; Conservative 3; Mismatches 13; Indels 0; Gaps 0;
Qy 7 HFAFFHQILATSTIAMSRCRTVYKQT 33
Db 60 HIPFFFIELCTLCMAIYSNPTVYST 86

```

RESULT 6
A86214
hypothetical protein [Imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
Change 03-Mar-2001 start change 17-May-2002

C: Accession: A86214
R: Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A: Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A: Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A: Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A: Reference number: A86141; MUID:21016719; PMID:11130712

A:Residues: 1-560 <STO>
A:Cross-references: GB:AE005172; NID:g8439883; PIDN:AAF75069.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: Arabidopsis hypothetical protein F13M22.23

| | | | | |
|-----------------------|--------------|----------------|----------------|-------------|
| Query Match | 26.4% | Score 53; | DB 2; | Length 560; |
| Best Local Similarity | 33.3% | Pred. NO. 9.9; | | |
| Matches 14; | Conservative | 5; | Mismatches 17; | Indels 6; |
| Gaps | 2; | | | |

QY 1 PYKWPFFHFAFFHQLLATSTI---AMSHCRP--TVYKOTSII 36
| | | | | : | | : | | : | | : |
Db 47 PLPWLFLEFLEFLLLIISTYGGGGRKTRCPSSSTYSIHSLSV 88

RESULT 7
C97662
hypothetical protein AGR_C_4580 [imported] - Agrobacterium tumefaciens (strain C58, C
C97662
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: C97662
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markeiz,
science 294, 2323-2328, 2001

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2003, 17:25:11 ; Search time 9 seconds
(without alignments)
165.905 Million cell updates/sec

Title: US-09-866-066-20

Perfect score: 201

Sequence: 1 PYKWPFFHFAFFHQLATSTIAMSHCRPTVYKQTSII 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 67 | 33.3 | 1001 | 1 | ORL1_DROME |
| 2 | 52 | 25.9 | 393 | 1 | SYDL_RAT |
| 3 | 51 | 25.4 | 355 | 1 | CRK3_MACMU |
| 4 | 49.5 | 24.6 | 213 | 1 | HLV3_BACSU |
| 5 | 49 | 24.4 | 555 | 1 | ALG9_YEAS |
| 6 | 49 | 24.4 | 1083 | 1 | EMBB_MYCLE |
| 7 | 48.5 | 24.1 | 485 | 1 | YGFO_ECOLI |
| 8 | 48 | 23.9 | 223 | 1 | ATP6_ONCMY |
| 9 | 48 | 23.9 | 335 | 1 | ARGC_PASNU |
| 10 | 48 | 23.9 | 449 | 1 | NTR_DROME |
| 11 | 47.5 | 23.6 | 437 | 1 | HGD_CAEEL |
| 12 | 47.5 | 23.6 | 1082 | 1 | EMBB_MYCSM |
| 13 | 47 | 23.4 | 2183 | 1 | RRPL_MEASA |
| 14 | 47 | 23.4 | 2183 | 1 | RRPL_MEASE |
| 15 | 46.5 | 23.1 | 350 | 1 | SYFA_THETH |
| 16 | 46 | 22.9 | 276 | 1 | PTP3_CHLEU |
| 17 | 46 | 22.9 | 355 | 1 | CRK3_CERAE |
| 18 | 46 | 22.9 | 394 | 1 | CIW3_HUMAN |
| 19 | 46 | 22.9 | 409 | 1 | CIW3_MOUSE |
| 20 | 46 | 22.9 | 411 | 1 | CIW3_RAT |
| 21 | 45.5 | 22.6 | 465 | 1 | YHJA_ECOLI |
| 22 | 45.5 | 22.6 | 1122 | 1 | TERT_MOUSE |
| 23 | 45 | 22.4 | 245 | 1 | CTRB_GADMO |
| 24 | 45 | 22.4 | 354 | 1 | CRK5_MOUSE |
| 25 | 45 | 22.4 | 365 | 1 | CIW9_CAVPO |
| 26 | 45 | 22.4 | 374 | 1 | CIW9_HUMAN |
| 27 | 45 | 22.4 | 462 | 1 | CATC_MOUSE |
| 28 | 45 | 22.4 | 572 | 1 | FDFT_USHMA |
| 29 | 45 | 22.4 | 1019 | 1 | VP3_RDVA |
| 30 | 45 | 22.4 | 1019 | 1 | VP3_RDVF |
| 31 | 45 | 22.4 | 1490 | 1 | CDR4_CANAL |
| 32 | 44.5 | 22.1 | 396 | 1 | ARSL_AQUAE |
| 33 | 44.5 | 22.1 | 453 | 1 | CARA_NEUCR |

RESULT 1

| ID | ORL1_DROME | STANDARD | PRT | 1001 AA |
|----|--|----------|-----|---------|
| AC | Q94526; | | | |
| DT | 16-OCT-2001 (Rel. 40, Created) | | | |
| DT | 16-OCT-2001 (Rel. 40, Last sequence update) | | | |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) | | | |
| DE | Open rectifier potassium channel protein 1 (Two pore domain potassium channel Ork1). | | | |
| GN | ORK1 OR CG1615. | | | |
| OS | Drosophila melanogaster (Fruit fly). | | | |
| OC | Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda; | | | |
| OC | Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; | | | |
| OC | Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. | | | |
| OX | NCBI_TaxID=7227; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Larva; | | | |
| RX | MEDLINE=97075152; PubMed=8917578; | | | |
| RA | Goldstein S.A.N., Price L.A., Rosenthal D.N., Pausch M.H.; | | | |
| RT | "ORK1, a potassium-selective leak channel with two pore domains | | | |
| RT | cloned from Drosophila melanogaster by expression in Saccharomyces | | | |
| RT | cerevisiae." | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 93:13256-13261(1996). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=Berkley; | | | |
| RX | MEDLINE=20196006; PubMed=10731132; | | | |
| RA | Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D., | | | |
| RA | Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F., | | | |
| RA | George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., | | | |
| RA | Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., | | | |
| RA | Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne C.R., Miklos G.L.G., | | | |
| RA | Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Pfeiffer B.D., | | | |
| RA | Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., | | | |
| RA | Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., | | | |
| RA | Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., | | | |
| RA | Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P., | | | |
| RA | Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., | | | |
| RA | Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P., | | | |
| RA | de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., | | | |
| RA | Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., | | | |
| RA | Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., | | | |
| RA | Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., | | | |
| RA | Glodek A., Gong F., Gorrell J.H., Gu H., Guan P., Harris M., | | | |
| RA | Harris N.I., Harvey D., Heiman T.J., Hernandez J.R., Houck J., | | | |
| RA | Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., | | | |
| RA | Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., | | | |
| RA | Kimmel B.E., Kodra C.D., Kraft C., Kravitz S., Kulp D., Lai Z., | | | |
| RA | Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., | | | |
| RA | Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D., | | | |
| RA | Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., | | | |
| RA | Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., | | | |
| RA | Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M., | | | |
| RA | Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., | | | |
| RA | Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., | | | |

ALIGNMENTS

| 34 | 44.5 | 22.1 | 493 | 1 | Y014_HUMAN | Q15048 homo sapien |
|----|------|------|------|---|------------|--------------------|
| 35 | 44.5 | 22.1 | 687 | 1 | CICL_HUMAN | P51800 homo sapien |
| 36 | 44.5 | 22.1 | 687 | 1 | CICL_HUMAN | P51801 homo sapien |
| 37 | 44.5 | 22.1 | 810 | 1 | PLMN_MACMU | P12545 macaca mula |
| 38 | 44.5 | 22.1 | 1337 | 1 | YDM5_SCHPO | P87136 schizosacch |
| 39 | 44 | 21.9 | 127 | 1 | YC03_MYCPN | Q50284 mycoplasma |
| 40 | 44 | 21.9 | 250 | 1 | DNK_DROME | Q36837 trichophyto |
| 41 | 44 | 21.9 | 269 | 1 | COX3_TRIRU | Q36837 trichophyto |
| 42 | 44 | 21.9 | 311 | 1 | YXXE_BACSU | Q07835 bacillus su |
| 43 | 44 | 21.9 | 355 | 1 | CKR3_HUMAN | P51677 homo sapien |
| 44 | 44 | 21.9 | 359 | 1 | CKR3_MOUSE | P51678 mus musculu |
| 45 | 44 | 21.9 | 382 | 1 | Y022_NPVOP | O10281 orgyia paeu |

RESULT 3
 CCR3_MACMU STANDARD; PRT; 355 AA.
 ID CCR3_MACMU
 AC P56483;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 3 (C-C CCR-3) (CCR-3) (CCR3)
 DE (CCR3)
 GN CCR3 OR CMKBR3.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21354176; PubMed-11461684;
 RA Margulies B.J., Hauer D.A., Clements J.E.;
 RA "Identification and comparison of eleven rhesus macaque chemokine
 RT receptors";
 RL AIDS Res. Hum. Retroviruses 17:981-986(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-98118446; PubMed-9454694;
 RA Soli N., Treboute C., Gomas E., Ferchal F., Shacklett B., Allison M.;
 RA "The rhesus macaque CCR3 chemokine receptor is a cell entry cofactor
 RT for HIV-2, but not for HIV-1";
 RL Virology 240:213-220(1998).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO EOTAXIN,
 CC MCP-3, MCP-4 AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: A017283; A070527.1;
 CC EMBL: Y13776; CA474107.1;
 CC InterPro: IPR000276; GPCR_Rhodopsin.
 CC Pfam: PF00001; 7tm_1; 1.
 CC PRINTS: PR00237; GPCR_Rhodopsin.
 CC PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; 1.
 CC PROSITE: PS00262; G-PROTEIN_RECEP_FL_2; 1.
 CC G-protein coupled receptor; Transmembrane.
 CC DOMAIN 1 34 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 35 62 1 (POTENTIAL).
 CC FT DOMAIN 63 72 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 73 93 2 (POTENTIAL).
 CC FT DOMAIN 94 107 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 108 129 3 (POTENTIAL).
 CC FT DOMAIN 130 146 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 147 171 4 (POTENTIAL).
 CC FT DOMAIN 172 203 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 204 223 5 (POTENTIAL).
 CC FT DOMAIN 224 239 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 240 264 6 (POTENTIAL).
 CC FT DOMAIN 265 281 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 282 305 7 (POTENTIAL).
 CC FT DOMAIN 306 355 CYTOPLASMIC (POTENTIAL).
 CC FT DISULFID 106 183 BY SIMILARITY.
 CC FT CONFLICT 180 183 K -> E (IN REF. 2).
 CC FT CONFLICT 202 202 K -> R (IN REF. 2).
 CC SEQUENCE 355 AA; 40805 MW; E271F1E694970D9F CRC64;
 CC
 CC Query Match 25.4%; Score 51; DB 1; Length 355;
 CC Best Local Similarity 46.2%; Pred. No. 6.5;

Matches 12; Conservative 2; Mismatches 8; Indels 4; Gaps 2;
 QY 7 HEAFFHOILATSTIAMSHC--RPTVY 30
 Db 278 HLDLF--VLATEVIAYSHCCVNPVIY 301
 RESULT 4
 HLX3_BACSU STANDARD; PRT; 213 AA.
 ID HLX3_BACSU
 AC P54175;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hemolysin III homolog.
 GN YPQ.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-168 / Marburg;
 RX MEDLINE-96349105; PubMed-8760912;
 RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
 RA Serror P.;
 RA "Sequence analysis of the Bacillus subtilis chromosome region between
 RT the serA and kag loci cloned in a yeast artificial chromosome.";
 RL Microbiology 142:2005-2016(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-168;
 RX MEDLINE-98044033; PubMed-9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Betero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriess R., Brunsch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
 RA Denizot K.D., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian F.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Hega K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Moone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Pressecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE UPF0073 (HLX-III) FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

FT TRANSMEM 525 542 POTENTIAL.
 FT TRANSMEM 555 572 POTENTIAL.
 FT TRANSMEM 576 598 POTENTIAL.
 FT TRANSMEM 611 633 POTENTIAL.
 FT TRANSMEM 648 710 POTENTIAL.
 FT TRANSMEM 690 712 POTENTIAL.
 SQ SEQUENCE 1083 AA; 117157 MW; 391009AF336F0DCF CRC64;

Query Match 24.48; Score 49; DB 1; Length 1083;
 Best Local Similarity 35.78; Pred. No. 36;
 Matches 10; Conservative 4; Mismatches 10; Indels 4; Gaps 1;

OY 1 PYKWPFFPAFF----HQIATSTAMSH 24
 ||||| : : : : :
 DB 573 PTKVHHEGLFAALGAAVAALTVLVSH 600
 ||||| : : : : :
 RESULT 7
 YGFO_ECOLI
 ID YGFO_ECOLI STANDARD; PRT; 485 AA.
 AC Q46815;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative purine permease ygfO.
 GN YGFO OR B2882 OR Z4221 OR ECS3755.
 OS Escherichia coli, and
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562, 83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE-21074935; PubMed-11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.P., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / RIMD 0509952;
 RX MEDLINE-21156231; PubMed-11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: BELONGS TO THE XANTHINE/URACIL PERMEASES FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL: U28375; AAA83063.1; -
 CC EMBL: AE000372; AAC75920.1; -
 CC EMBL: AE000518; AAG58011.1; -
 CC EMBL: AF002563; BAB37178.1; -
 CC DR EcoGene; EGI3085; ygfO.
 CC DR InterPro; IPR000444; Xan_ur_permease.
 CC DR Pfam; PF00860; xan_ur_permease; 1.
 CC DR TIGRfam; TIGR00801; ncs2; 1.
 CC DR PROSITE; PS01116; XANTH_URACIL_PERMASE; 1.
 CC KW Hypothetical protein; Transmembrane; transport; Complete proteome.
 CC FT TRANSMEM 45 65 POTENTIAL.
 CC FT TRANSMEM 75 95 POTENTIAL.
 CC FT TRANSMEM 100 120 POTENTIAL.
 CC FT TRANSMEM 140 160 POTENTIAL.
 CC FT TRANSMEM 171 191 POTENTIAL.
 CC FT TRANSMEM 200 220 POTENTIAL.
 CC FT TRANSMEM 230 250 POTENTIAL.
 CC FT TRANSMEM 318 338 POTENTIAL.
 CC FT TRANSMEM 362 382 POTENTIAL.
 CC FT TRANSMEM 383 403 POTENTIAL.
 CC FT TRANSMEM 416 436 POTENTIAL.
 CC FT TRANSMEM 445 465 POTENTIAL.
 CC SQ SEQUENCE 485 AA; 51361 MW; 07EFDCC3795AC636 CRC64;
 Query Match 24.1%; Score 48.5; DB 1; Length 485;
 Best Local Similarity 50.0%; Pred. No. 20;
 Matches 10; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

OY 1 PYKWPFFPAFFHQIATSTI 20
 ||| : : : : :
 DB 266 PFKYGFSGFS-FHQELVVGTI 284
 ||| : : : : :
 RESULT 8
 ATP6_ONCMY
 ID ATP6_ONCMY STANDARD; PRT; 223 AA.
 AC P48178;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ATP synthase A chain (EC 3.6.3.14) (Protein 6).
 GN MTATP6 OR ATP6.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OG Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Liver.
 RX MEDLINE-96139027; PubMed-8587139;
 RA Zardoya R., Garrido-Perterra A., Bautista J.M.;
 RA "The complete nucleotide sequence of the mitochondrial DNA genome of
 the rainbow trout, *Oncorhynchus mykiss*.";
 RL J. Mol. Evol. 41:942-951(1995).
 CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A
 CC DIRECT ROLE IN THE TRANSLLOCATION OF PROTONS ACROSS THE MEMBRANE.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) -> ADP + phosphate +
 CC H(+) (out).
 CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
 CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
 CC HAS THREE MAIN SUBUNITS: A, B AND C.
 CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

QY 1 PYKWPFFFAFFHOILATSTIAMSHCRPTVY 30
 DB 347 PYWV---FAP-----HWLAMSHCYCNPIIY 368

RESULT 11
 HGD_CABEL STANDARD; PRT; 437 AA.

AC Q9Y041; G62087; Q9NJP3;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Homogentisate 1,2-dioxygenase (EC 1.13.11.5) (Homogentisicase)
 DE Homogentisate oxygenase (Homogentisic acid oxidase).
 GN HGO-1 OR W05D4.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Schmidt S.R., Werner E., Mueller C.R., Kress W.;
 RT "Cloning and characterization of the homogentisate 1,2-dioxygenase
 gene in A. thaliana and C. elegans.";
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2; TISSUE-Embryo;
 RA Schmidt S.R., Werner E., Mueller C.R., Kress W.;
 RT "Sequence homology of HGO genes in eukaryotic organisms.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Kershaw J.;
 RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP REVISIONS.
 RA Durbán R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Homogentisate + O(2) -> 4-maleylacetoacetate.
 CC -1- COFACTOR: IRON.
 CC -1- PATHWAY: Catabolism of tyrosine; third step.
 CC -1- PATHWAY: Catabolism of phenylalanine; fourth step.
 CC -1- SIMILARITY: BELONGS TO THE HOMOGENTISATE DIOXYGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF136150; AAF61419.1;
 DR EMBL; U95181; AAD00776.1;
 DR EMBL; Z93778; CAB07848.4;
 DR EMBL; AL034364; CAB07848.4; JOINED.
 DR EMBL; AL034364; CAB07848.4; JOINED.
 DR EMBL; Z93778; CAA22255.4;
 DR HSSP; Q93099; 1BYB.
 DR WormPep; W06D4.1; CE29602.
 DR TIGRFAms; TIGR01015; hnga.1.
 KW Oxidoreductase; Dioxygenase; Iron; Phenylalanine catabolism;
 FT Tyrosine catabolism.
 FT METAL 336 336 IRON (BY SIMILARITY).
 FT METAL 342 342 IRON (BY SIMILARITY).
 FT METAL 372 372 IRON (BY SIMILARITY).
 FT CONFLICT 156 156 D -> N (IN REF. 2).
 FT CONFLICT 176 176 L -> P (IN REF. 2).
 FT CONFLICT 193 193 V -> G (IN REF. 2).

FT CONFLICT 242 242 V -> G (IN REF. 2).
 FT CONFLICT 265 265 D -> E (IN REF. 2).
 FT CONFLICT 433 433 F -> Y (IN REF. 2).
 SQ SEQUENCE 437 AA; 49238 MW; C14E7077C7CF9703 CRC64;

Query Match 23.6%; Score 47.5; DB 1; Length 437;
 Best Local Similarity 23.3%; Pred. No. 25;
 Matches 7; Conservative 9; Mismatches 9; Indels 5; Gaps 1;

QY 1 PYKWPFFFAFFHOILATSTIAMSHCRPTVY 30
 DB 275 PYKD-----LKKFMVINTVSPDCHDPSIF 299

RESULT 12
 EMBL_MYCSM STANDARD; PRT; 1082 AA.

AC Q50395; O30406;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable arabinosyltransferase B (EC 2.4.2.-).
 GN EMBL.
 OS Mycobacterium smegmatis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1772;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-imm30;
 RA Telenti A., Philipp W., Sreevatsan S., Bernasconi C., Stockbauer K.E.,
 RT Wiele B., Musser J.M., Jacobs W.R.;
 RT "The emb operon, a gene cluster of Mycobacterium tuberculosis involved
 in resistance to ethambutol.";
 RN Nat. Med. 3:567-570(1997).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANTS EMB RESISTANT F-289; M-289 AND T-292.
 RC STRAIN-imm30;
 RX MEDLINE=98080198; PubMed=9420031;
 RA Lety M.A., Nair S., Berche P., Escuyer V.;
 RT "A single point mutation in the embB gene is responsible for
 resistance to ethambutol in Mycobacterium smegmatis.";
 RL Antimicrob. Agents Chemother. 41:2629-2633(1997).
 CC -1- FUNCTION: Arabinosyl transferase responsible for the
 CC polymerization of arabinose into the arabinan of arabinogalactan.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- MISCELLANEOUS: This is one of the target of the anti-tuberculosis
 CC drug ethambutol [(S,S')-2,2'-(ethylenedimino)di-1-butanol; EMB].
 CC EMB is a first-line drug used to treat tuberculosis. EMB inhibits
 CC the transfer of arabinogalactan into the cell wall.
 CC -1- SIMILARITY: BELONGS TO THE EMB FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U46844; AAC45273.1;
 DR EMBL; AF004289; AAB69157.1;
 DR EMBL; AF004289; AAB69157.1;
 KW Transferase; Glycosyltransferase; Transmembrane; Cell wall;
 FT Antibiotic resistance.
 FT TRANSMEM 28 50 POTENTIAL.
 FT TRANSMEM 223 241 POTENTIAL.
 FT TRANSMEM 262 281 POTENTIAL.
 FT TRANSMEM 333 352 POTENTIAL.
 FT TRANSMEM 359 381 POTENTIAL.
 FT TRANSMEM 420 442 POTENTIAL.
 FT TRANSMEM 462 481 POTENTIAL.
 FT TRANSMEM 522 544 POTENTIAL.

01-AUG-1992 (Rel. 23, Last sequence update)
 16-OCT-2001 (Rel. 40, Last annotation update)
 Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20) (Phenylalanine-tRNA ligase alpha chain) (PHERS).
 PHERS.
 Thermus thermophilus.
 Bacteria; Thermus/deinococcus group; Deinococci; Thermales;
 OC Thermaceae; Thermus.
 OX NCBI_TaxID=274;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-HB8 / ATCC 27634;
 RX MEDLINE-93083630; PubMed-1451792;
 RA Keller B., Kast P., Hennecke H.;
 RT "Cloning and sequence analysis of the phenylalanyl-tRNA synthetase
 genes (pheST) from Thermus thermophilus.";
 RL FEBS Lett. 301:83-88(1992).
 RN [2]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-29.
 RC STRAIN-HB8 / ATCC 27634;
 RX MEDLINE-92375722; PubMed-1508711;
 RA Kreutzer R., Kluft V., Bobkova E.V., Lavrik O.J., Sprinzl M.;
 RT "Structure of the phenylalanyl-tRNA synthetase genes from Thermus
 thermophilus HB8 and their expression in Escherichia coli.";
 RL Nucleic Acids Res. 20:4173-4178(1992).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN-HB8 / ATCC 27634;
 RA Lechler A., Kreutzer R.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RC STRAIN-HB8 / ATCC 27634;
 RX MEDLINE-94257735; PubMed-8199244;
 RA Mosyak L., Safto M.;
 RT "Phenylalanyl-tRNA synthetase from Thermus thermophilus has four
 antiparallel folds of which only two are catalytically functional.";
 RL Biochimie 75:1091-1098(1993).
 RN [5]
 RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RC STRAIN-HB8 / ATCC 27634;
 RX MEDLINE-95333238; PubMed-7664121;
 RA Mosyak L., Reshetnikova L., Goldgur Y., Delarue M., Safto M.G.;
 RT "Structure of phenylalanyl-tRNA synthetase from Thermus
 thermophilus.";
 RL Nat. Struct. Biol. 2:537-547(1995).
 RN [6]
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RC MEDLINE-99194858; PubMed-10092459;
 RA Reshetnikova L., Moor N., Lavrik O., Vassilyev D.G.;
 RT "Crystal structures of phenylalanyl-tRNA synthetase complexed with
 phenylalanine and a phenylalanyl-adenylate analogue.";
 RL J. Mol. Biol. 287:555-568(1999).
 CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) -> AMP +
 CC diphosphate + L-phenylalanyl-tRNA(Phe).
 CC -1- SUBUNIT: Tetramer of two alpha and two beta chains.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC PHE-TRNA SYNTHETASE ALPHA CHAIN SUBFAMILY 1.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: Z12118; CAA78104.1; -;
 CC EMBL: X65609; CAA6559.1; -;
 CC EMBL: Y15464; CAA75644.1; -;
 CC PIR: S21468; S21468.
 CC PIR: S22366; S22366.

DR PIR: S25717; S25717.
 DR PDB: 1PYS; 19-NOV-97.
 DR PDB: 1B70; 09-FEB-00.
 DR PDB: 1B7Y; 26-JAN-00.
 DR InterPro: IPR002106; AaTrnaLigaseII.
 DR InterPro: IPR004529; PheS.
 DR InterPro: IPR004188; Phe_tRNA_synt_N.
 DR InterPro: IPR002319; tRNA_synt_2d.
 DR Pfam: PF01409; tRNA_synt_2d; 1.
 DR Pfam: PF02912; Phe_tRNA_synt_N; 1.
 DR TIGRFAMS: TIGR00468; PheS; 1.
 DR PROSITE: PS50862; AA_TRNA_LIGASE_II; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW 3D-structure.
 SQ SEQUENCE 350 AA; 39259 MW; DEE0237F7CD9A461 CRC64;
 Query Match 23.1%; Score 46.5; DB 1; Length 350;
 Best Local Similarity 35.7%; Pred. No. 28;
 Matches 10; Conservative 6; Mismatches 9; Indels 3; Gaps 1;
 QY 7 HFAFFHQ---ILATSTIANSHCRPTVVK 31
 | | | | | : : | | | : | :
 Db 212 HEAVFHOLEGLVVGEGIAHAHLKGAIVE 239
 Search completed: March 24, 2003, 17:26:13
 Job time : 12 secs

```

FT TRANSMEM 557 574 POTENTIAL.
FT TRANSMEM 578 600 POTENTIAL.
FT TRANSMEM 613 635 POTENTIAL.
FT TRANSMEM 650 672 POTENTIAL.
FT TRANSMEM 689 711 POTENTIAL.
FT VARIANT 289 289 I -> M (RESISTANCE TO EMB).
FT VARIANT 289 289 I -> F (RESISTANCE TO EMB).
FT VARIANT 292 292 M -> T (RESISTANCE TO EMB).
FT CONFLICT 453 453 I -> M (IN REF. 2).
FT CONFLICT 459 459 D -> V (IN REF. 2).
FT CONFLICT 508 508 K -> N (IN REF. 2).
FT CONFLICT 579 579 T -> I (IN REF. 2).
SQ SEQUENCE 1082 AA; 116776 MW; 10883EA54F72E0FE CRC64;

Query Match 23.6%; Score 47.5; DB 1; Length 1082;
Best Local Similarity 34.3%; Pred. No. 59;
Matches 12; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

QY 1 PYKWPFFHFAFFHOI-----LATSTIAMSHCRPTVYK 31
| | | | | | | | | | | | | | | | | | | | |
DB 575 PKWTHHFGFLGAAGVGAAMALATVLVS---PTVLR 606
| | | | | | | | | | | | | | | | | | | | |

RESULT 13
RRPL_MEASA
ID RRPL_MEASA STANDARD; PRT; 2183 AA.
AC P35975;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA polymerase beta subunit (EC 2.7.7.48) (Large structural protein)
DE (L protein).
GN L.
OS Measles virus (strain Aik-C) (Subacute sclerose panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=36408;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227570; PubMed=8470368;
RA Mori T., Sasaki K., Hashimoto H., Makino S.;
RT Molecular cloning and complete nucleotide sequence of genomic RNA of
RT the Aik-C strain of attenuated measles virus.;
RL Virus Genes 7:67-81(1993).
CC -1- FUNCTION: PROBABLE COMPONENT OF THE ACTIVE POLYMERASE. IT MAY
CC SYNTHESIS OF NEWLY SYNTHESIZED VIRAL MRNAs, RNA EDITING OF THE P
CC GENE TRANSCRIPT, AND PROTEIN KINASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SIMILARITY: WITH L PROTEIN OF OTHER PARAMYXOVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S58435; AAB26147.1; -.
CC PIR; G48556; G48556.
CC InterPro; IPR001016; Viral_RNA_pol_L.
CC Pfam; PF00946; Paramyx_RNA_pol; 1.
CC Transferase; RNA-directed RNA polymerase.
CC SEQUENCE 2183 AA; 247754 MW; 12A6686F4CA62199 CRC64;

Query Match 23.4%; Score 47; DB 1; Length 2183;
Best Local Similarity 39.3%; Pred. No. 1.4e+02;
Matches 11; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 2 YKWPFFHFAFFHOIATSTIAMS--HCRP 27
| | | | | | | | | | | | | | | | | | | | |
DB 687 YGLPSFFQWLHKLRLTSVLVSDPHCPP 714
| | | | | | | | | | | | | | | | | | | | |

RESULT 15
SYFA_THETH
ID SYFA_THETH STANDARD; PRT; 350 AA.
AC P27001;
DT 01-AUG-1992 (Rel. 23, Created)

```

```

DB 687 YGLPSFFQWLHKLRLTSVLVSDPHCPP 714
| | | | | | | | | | | | | | | | | | | | |

RESULT 14
RRPL_MEASE
ID RRPL_MEASE STANDARD; PRT; 2183 AA.
AC P12576;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA polymerase beta subunit (EC 2.7.7.48) (Large structural protein)
DE (L protein).
GN L.
OS Measles virus (strain Edmonston) (Subacute sclerose panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11235;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219537; PubMed=2835864;
RA Blumberg B.M., Crowley J.C., Silverman J.I., Menonna J., Cook S.D.,
RA Dowling P.C.;
RT Measles virus L protein evidences elements of ancestral RNA
RT polymerase.;
RL Virology 164:487-497(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90085790; PubMed=2596022;
RA Cattaneo R., Schmid A., Spielhofer P., Kaelin K., Baczko K.,
RA Meulen V., Pardowitz J., Flanagan S., Rima B.K., Udem S.A.;
RT Mutated and hypermutated genes of persistent measles viruses which
RT caused lethal human brain diseases.;
RL Virology 173:415-425(1989).
CC -1- FUNCTION: PROBABLE COMPONENT OF THE ACTIVE POLYMERASE. IT MAY
CC SYNTHESIS OF NEWLY SYNTHESIZED VIRAL MRNAs, RNA EDITING OF THE P
CC GENE TRANSCRIPT, AND PROTEIN KINASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SIMILARITY: WITH L PROTEIN OF OTHER PARAMYXOVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M20865; AAA46430.1; -.
CC EMBL; K01711; AAA75501.1; -.
CC PIR; A28919; ZLNZMV.
CC InterPro; IPR001016; Viral_RNA_pol_L.
CC Pfam; PF00946; Paramyx_RNA_pol; 1.
CC Transferase; RNA-directed RNA polymerase.
CC SEQUENCE 2183 AA; 247646 MW; 1B0B03CA2E2B6EA5 CRC64;

Query Match 23.4%; Score 47; DB 1; Length 2183;
Best Local Similarity 39.3%; Pred. No. 1.4e+02;
Matches 11; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 2 YKWPFFHFAFFHOIATSTIAMS--HCRP 27
| | | | | | | | | | | | | | | | | | | | |
DB 687 YGLPSFFQWLHKLRLTSVLVSDPHCPP 714
| | | | | | | | | | | | | | | | | | | | |

RESULT 15
SYFA_THETH
ID SYFA_THETH STANDARD; PRT; 350 AA.
AC P27001;
DT 01-AUG-1992 (Rel. 23, Created)

```



```

CC EMBL: L77246; AAA96637.1; -.
DR EMBL: Z99115; CAB14097.1; -.
DR SUBMILet: BG11620; YP10.
DR InterPro: IPR004254; HlyIII_related.
DR Pfam: PF03006; UPF0073; 1.
DR TIGRfam: TIGR01065; hlyIII; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 41 61 POTENTIAL.
FT TRANSMEM 75 95 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
FT TRANSMEM 127 147 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
SQ SEQUENCE 213 AA; 23588 MW; 89DD028CB7EAB2AE CRC64;

Query Match 24.6%; Score 49.5; DB 1; Length 213;
Best Local Similarity 37.0%; Pred. No. 6.5;
Matches 10; Conservative 3; Mismatches 11; Indels 3; Gaps 1;

OY 2 YKW---PFHFAFFHOILATSTIAMSHC 25
| | | | | : | | |
DB 175 YINKPPEHHAHWSFVLGSAAMFPC 201

RESULT 5
ALG9_YEAST STANDARD; PRT; 555 AA.
AC P53868;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 13-JUN-2002 (Rel. 41, Last annotation update)
DE Probable mannoseyltransferase ALG9 (EC 2.4.1.-).
GN ALG9 OR YNL219C OR N1295.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID:4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:SS328;
RX MEDLINE-96293493; PubMed-8692962;
RA Burda P., Te Heesen S., Brachat A., Wach A., Duesterhoeft A., Aebi M.;
RT "Stepwise assembly of the lipid-linked oligosaccharide in the
endoplasmic reticulum of Saccharomyces cerevisiae: identification of
the ALG9 gene encoding a putative mannoseyl transferase.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:7160-7165(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Duesterhoeft A., Floeth M., Fritz C., Heuss-Neitzel D.,
RA Hilbert H., Moestl D.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE TRANSFER OF MANNOSE FROM DOL-P-MAN TO
LIPID-LINKED OLIGOSACCHARIDES.
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
reticulum (By similarity).
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 22.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: X96417; CAA65277.1; -.
DR EMBL: Z71495; CAA96122.1; -.
DR SGD: S0005163; ALG9.
KW Transferase; Glycosyltransferase; Transmembrane;
Endoplasmic reticulum.

```

```

FT TRANSMEM 8 28 POTENTIAL.
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 114 134 POTENTIAL.
FT TRANSMEM 177 197 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 269 289 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 350 370 POTENTIAL.
SQ SEQUENCE 555 AA; 63776 MW; 108ED4E0B0C2AAA7 CRC64;

Query Match 24.4%; Score 49; DB 1; Length 555;
Best Local Similarity 31.2%; Pred. No. 19;
Matches 10; Conservative 4; Mismatches 8; Indels 10; Gaps 1;

OY 4 WPF-----HFPAFFHOILATSTIAMSHC 25
| | | | | : | | | | |
DB 184 WPFVLILSLPLCLHLYLFNHRIRIISTIRTAFDCC 215

RESULT 6
EMBB_MYCLE STANDARD; PRT; 1083 AA.
AC Q9CDA9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable arabinosyltransferase B (EC 2.4.2.-).
GN EMBB OR M0104.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID:1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:TN;
RX MEDLINE-21128732; PubMed-11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- FUNCTION: Arabinosyl transferase responsible for the
polymerization of arabinose into the arabinan of arabinogalactan
(By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO THE EMB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: AL583917; CAC29612.1; -.
DR Leproma; M0104; -.
KW Transferase; Glycosyltransferase; Transmembrane; Cell wall;
Complete proteome.
FT TRANSMEM 23 45 POTENTIAL.
FT TRANSMEM 222 239 POTENTIAL.
FT TRANSMEM 252 274 POTENTIAL.
FT TRANSMEM 331 350 POTENTIAL.
FT TRANSMEM 357 379 POTENTIAL.
FT TRANSMEM 421 443 POTENTIAL.
FT TRANSMEM 456 478 POTENTIAL.

```

| | Query Match | 25.9% | Score 52; | DB 1; | Length 393; |
|-----------------------|-------------|---------------------------------|----------------|------------|-------------|
| Best Local Similarity | 43.3% | | Pred. NO. 5.2; | | |
| Matches | 13; | Conservative | 4; | Mismatches | 11; |
| | | | | Indels | 2; |
| | | | | Gaps | 1; |
| QY | 9 | APFH--QLLATSTIAMSHCRPTVYKOTSII | 36 | | |
| | | | : | | |
| | | : | : | | |
| DB | 14 | APFHLCCTLIAGQHLGKMTKCNITCKHWTSP | 43 | | |

| Result No. | Query | | | DB | ID | Description |
|------------|-------|-------|--------|----|---------------------|--------------------|
| | Score | Match | Length | | | |
| 1 | 67 | 33.3 | 618 | 1 | US-08-333-3112-2 | Sequence 2, Appli |
| 2 | 47.5 | 23.6 | 391 | 4 | US-09-134-001C-3442 | Sequence 3442, Ap |
| 3 | 47.5 | 23.6 | 3165 | 2 | US-08-459-146-3 | Sequence 3, Appli |
| 4 | 47.5 | 23.6 | 3165 | 2 | US-08-459-065-3 | Sequence 3, Appli |
| 5 | 47 | 23.4 | 535 | 3 | US-08-369-822C-24 | Sequence 24, Appli |
| 6 | 47 | 23.4 | 535 | 3 | US-08-583-776C-39 | Sequence 39, Appli |
| 7 | 47 | 23.4 | 535 | 3 | US-08-434-831B-36 | Sequence 36, Appli |
| 8 | 47 | 23.4 | 2183 | 1 | US-08-348-831A-7 | Sequence 7, Appli |
| 9 | 47 | 23.4 | 2183 | 2 | US-08-905-817-7 | Sequence 7, Appli |
| 10 | 46 | 22.9 | 394 | 4 | US-09-144-914-4 | Sequence 4, Appli |
| 11 | 46 | 22.9 | 405 | 4 | US-09-144-914-5 | Sequence 5, Appli |
| 12 | 45 | 22.4 | 197 | 1 | US-08-456-840-48 | Sequence 48, Appli |
| 13 | 45 | 22.4 | 197 | 1 | US-08-266-407A-48 | Sequence 48, Appli |
| 14 | 45 | 22.4 | 197 | 2 | US-08-892-544-48 | Sequence 48, Appli |
| 15 | 45 | 22.4 | 260 | 4 | US-09-134-001C-4507 | Sequence 4507, Ap |
| 16 | 45 | 22.4 | 354 | 4 | US-08-724-984A-2 | Sequence 2, Appli |
| 17 | 45 | 22.4 | 944 | 4 | US-09-134-001C-4352 | Sequence 4352, Ap |
| 18 | 45 | 22.4 | 1091 | 3 | US-08-633-768A-2 | Sequence 2, Appli |
| 19 | 44.5 | 22.1 | 278 | 1 | US-08-392-828C-4 | Sequence 4, Appli |
| 20 | 44.5 | 22.1 | 278 | 3 | US-09-330-945-4 | Sequence 4, Appli |
| 21 | 44 | 21.9 | 300 | 1 | US-07-640-039-5 | Sequence 5, Appli |
| 22 | 44 | 21.9 | 300 | 4 | US-08-439-932A-5 | Sequence 5, Appli |
| 23 | 44 | 21.9 | 302 | 1 | US-07-640-029-6 | Sequence 6, Appli |
| 24 | 44 | 21.9 | 302 | 1 | US-07-921-807B-7 | Sequence 7, Appli |
| 25 | 44 | 21.9 | 302 | 1 | US-07-921-807B-8 | Sequence 8, Appli |
| 26 | 44 | 21.9 | 302 | 1 | US-08-441-944A-7 | Sequence 7, Appli |
| 27 | 44 | 21.9 | 302 | 1 | US-08-441-944A-8 | Sequence 8, Appli |

RESULT 5

US-08-369-822C-24
; Sequence 24, Application US/08369822C
; Patent No. 6015660
; GENERAL INFORMATION:
; APPLICANT: Lipkin, W. I.
; APPLICANT: Briese, Thomas
; APPLICANT: Kliche, Stefanie
; APPLICANT: Schneider, Patrick A.
; APPLICANT: Stitz, Lothar
; APPLICANT: Schneemann, Anette
; TITLE OF INVENTION: Borna Disease Viral Sequences,
; TITLE OF INVENTION: Diagnostics and Therapeutics for Central Nervous
; TITLE OF INVENTION: System Diseases
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 865 South Figueroa Street, 29th Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90017-2571
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS NT-WORDPERFECT 8.0
; SOFTWARE: ASCII (DOS) TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/369,822C
; FILING DATE: 06-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Churchill, Margaret A. (Ph.D.)
; REGISTRATION NUMBER: 39,944
; REFERENCE/DOCKET NUMBER: 1279-194XX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213/892-9200
; TELEFAX: 213/680-4518
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 535 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-369-822C-24

Query Match 23.4%; Score 47; DB 3; Length 535;
Best Local Similarity 39.3%; Pred. No. 50;
Matches 11; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 2 YKPFHFAFFHQLATSTIAMS--HCRP 27

Db 259 YGLPSFFQWLKRLKTSVLYSDPHCPP 286

RESULT 6

US-08-582-776C-39
; Sequence 39, Application US/08582776C
; Patent No. 6077510
; GENERAL INFORMATION:
; APPLICANT: Lipkin, W. I.
; APPLICANT: Briese, Thomas
; APPLICANT: Kliche, Stefanie
; APPLICANT: Schneider, Patrick A.
; APPLICANT: Stitz, Lothar
; APPLICANT: Schneemann, Anette
; TITLE OF INVENTION: Borna Disease Viral Sequences,
; TITLE OF INVENTION: Diagnostics and Therapeutics for Central Nervous
; TITLE OF INVENTION: System Diseases
; NUMBER OF SEQUENCES: 61

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 865 South Figueroa Street, 29th Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90017-2576
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS NT
; SOFTWARE: ASCII DOS TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/582,776C
; FILING DATE: 04-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/369,822
; FILING DATE: 06-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/434,831
; FILING DATE: 04-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Churchill, Margaret A.
; REGISTRATION NUMBER: 39,944
; REFERENCE/DOCKET NUMBER: 1279-194C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213/892-9200
; TELEFAX: 213/680-4518
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 535 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-582-776C-39

Query Match 23.4%; Score 47; DB 3; Length 535;
Best Local Similarity 39.3%; Pred. No. 50;
Matches 11; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 2 YKPFHFAFFHQLATSTIAMS--HCRP 27

Db 259 YGLPSFFQWLKRLKTSVLYSDPHCPP 286

RESULT 7

US-08-434-831B-36
; Sequence 36, Application US/08434831B
; Patent No. 6113905
; GENERAL INFORMATION:
; APPLICANT: Lipkin, W. I.
; APPLICANT: Briese, Thomas
; APPLICANT: Kliche, Stefanie
; APPLICANT: Schneider, Patrick A.
; APPLICANT: Stitz, Lothar
; APPLICANT: Schneemann, Anette
; TITLE OF INVENTION: Borna Disease Viral Sequences,
; TITLE OF INVENTION: Diagnostics and Therapeutics for Central Nervous
; TITLE OF INVENTION: System Diseases
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 865 South Figueroa Street, 29th Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90017-2571
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

; TYPE: PRT
; ORGANISM: Murine
; FEATURE:
; OTHER INFORMATION: TASK
US-09-144-914-5

Query Match      22.9%; Score 46; DB 4; Length 405;
Best Local Similarity 31.2%; Pred. No. 51;
Matches 10; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

QY    1 PYKWF---WPFHFAFFHQILATSTIAMSHCRPT 28
      ||| | :| :| :|||
DB    68 PHKAGVQWRFGSGFYFAYIVITTYIGYGHRAPS 39

RESULT 12
US-08-456-840-48
; Sequence 48, Application US/08456840
; Patent No. 5597908
; GENERAL INFORMATION:
; APPLICANT: Taddei-Peters, W. C.
; APPLICANT: Butler, Sandra M.
; TITLE OF INVENTION: Immunoactive Peptides of Apo(a)
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo NO. 5597908el
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/266,407
; FILING DATE: 27-JUN-1994
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
US-08-456-840-48

Query Match      22.4%; Score 45; DB 1; Length 197;
Best Local Similarity 27.9%; Pred. No. 31;
Matches 12; Conservative 5; Mismatches 14; Indels 12; Gaps 2;

QY    1 PYKWPF-----HFAPFHQILATSTIAMSEC---RPTVK 31
      ||| | :| :| :|||
DB    8 PHSWPQVSLTRTFGCGGLTISPENVLTAACLKSRPSSYK 50

RESULT 13
US-08-266-407A-48
; Sequence 48, Application US/08266407A

```

```

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/266.407
;   FILING DATE: 27-JUN-1994
;   APPLICATION NUMBER: US 08/172.461
;   FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
;   NAME: Gormley, Mary E.
;   REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 301-258-5200
;   TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; US-08-892-544-48

Query Match                22.4%; Score 45; DB 2; Length 197;
Best Local Similarity      27.9%; Pred No. 31;
Matches 12; Conservative 5; Mismatches 14; Indels 12; Gaps 2;

QY 1 PYKWPF-----HFAFFHQILATSTIAMSHC---RPTVTK 31
    ||: ||: | : || ||: ||
Db 8 PHSWPQVSLRTFGHGCGTGLISPEWLTAAHLCKLSRPSSVK 50

RESULT 15
US-09-134-001C-4507
; Sequence 4507, Application US/091344001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4507
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4507

Query Match                22.4%; Score 45; DB 4; Length 260;
Best Local Similarity      28.6%; Pred. No. 43;
Matches 6; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 10 FFHQILATSTIAMSHCRPTYV 30
    ||: : : : ||| ::|
Db 179 YFHCLSHLLTLTHCRMXY 199

Search completed: March 24, 2003, 17:28:04
Job time : 16 secs
```

OPERATING SYSTEM: WINDOWS NT-WORDPERFECT 8.0

SOFTWARE: ASCII (DOS) TEXT

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/434,831B

FILING DATE: 04-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA: US 08/369,822

FILING DATE: 06-JAN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Churchill, Margaret A. (Ph.D.)

REGISTRATION NUMBER: 39,944

REFERENCE/DOCKET NUMBER: 1279-194C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 213/892-9200

TELEFAX: 213/680-4518

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 535 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-434-831B-36

Query Match 23.4%; Score 47; DB 3; Length 535;

Best Local Similarity 39.3%; Pred. No. 50;

Matches 11; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 2 YKWPFFAFPHQILATSIAMS--HCRP 27

Db 259 YGLPFFQWLKRLTSVLYVSDPHCPP 286

RESULT 8

US-08-348-891A-7

; Sequence 7, Application US/08348891A

; Patent No. 5654136

; GENERAL INFORMATION:

; APPLICANT: SASAKI, Keiko

; APPLICANT: MORI, Takayuki

; APPLICANT: MAKINO, Satoshi

; TITLE OF INVENTION: ATTENUATED MEASLES VIRUS VACCINE,

; TITLE OF INVENTION: CONTAINING SPECIFIC NUCLEOTIDE SEQUENCE AND A METHOD FOR

; TITLE OF INVENTION: ITS ABSOLUTE IDENTIFICATION

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: YOUNG & THOMPSON

; STREET: 745 South 23rd Street

; CITY: Arlington

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA: US/08/348,891A

; APPLICATION NUMBER: 32,925

; FILING DATE: 25-NOV-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/848,400

; FILING DATE: 10-MAR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 3-293625

; FILING DATE: 14-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: PATCH, Andrew J.

; REGISTRATION NUMBER: 32,925

; REFERENCE/DOCKET NUMBER: KP-7501

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-521-2297

TELEFAX: 703-685-0573

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 2183 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-348-891A-7

Query Match 23.4%; Score 47; DB 1; Length 2183;

Best Local Similarity 39.3%; Pred. No. 2.5e+02;

Matches 11; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 2 YKWPFFAFPHQILATSIAMS--HCRP 27

Db 687 YGLPFFQWLKRLTSVLYVSDPHCPP 714

RESULT 9

US-08-905-817-7

; Sequence 7, Application US/08905817

; Patent No. 5824777

; GENERAL INFORMATION:

; APPLICANT: SASAKI, Keiko

; APPLICANT: MORI, Takayuki

; APPLICANT: MAKINO, Satoshi

; TITLE OF INVENTION: ATTENUATED MEASLES VIRUS VACCINE,

; TITLE OF INVENTION: CONTAINING SPECIFIC NUCLEOTIDE SEQUENCE AND A METHOD FOR

; TITLE OF INVENTION: ITS ABSOLUTE IDENTIFICATION

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: YOUNG & THOMPSON

; STREET: 745 South 23rd Street

; CITY: Arlington

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/905,817

; FILING DATE: 04-AUG-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/348,891

; FILING DATE: 25-NOV-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/848,400

; FILING DATE: 10-MAR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 3-293625

; FILING DATE: 14-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: PATCH, Andrew J.

; REGISTRATION NUMBER: 32,925

; REFERENCE/DOCKET NUMBER: KP-7501A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-521-2297

; TELEFAX: 703-685-0573

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2183 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-905-817-7

RESULT 2
US-09-134-001C-3442
; Sequence 3442, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3442
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3442

Query Match 23.6%; Score 47.5; DB 4; Length 391;
Best Local Similarity 35.3%; Pred. No. 29;
Matches 12; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

QY 2 YKWPFFAFHHQILATSTIAMSHCRPTVYKQTSI 35
| | | | | : | : | : | | | | |
DB 145 YKWPFFAPDQH-IETVNTLTIPDRVPVFKLSM 177

RESULT 3
US-08-459-065-3
; Sequence 3, Application US/08459146
; Patent No. 5866405
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil Ho
; APPLICANT: Nuss, Donald Lee
; TITLE OF INVENTION: Genetically Engineered Transmissible
; TITLE OF INVENTION: Hypovirulence
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,146
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/832,117
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Catherine R
; REGISTRATION NUMBER: 34,240
; REFERENCE/DOCKET NUMBER: 8589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-6208
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3165 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Endothia parasitica (Cryptonectria
; ORGANISM: Parasitica)
; STRAIN: EP713
; US-08-459-065-3

Query Match 23.6%; Score 47.5; DB 2; Length 3165;
Best Local Similarity 34.4%; Pred. No. 3.2e+02;
Matches 11; Conservative 8; Mismatches 8; Indels 5; Gaps 1;

QY 3 KWPFFAFHHQILATSTIAMSHCRPTVYKQTS 34
| | | | | : | : | : | | | | |
DB 1991 KYPAHFATFSELLAHNNVNV-----TEWKRLS 2017

; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Endothia parasitica (Cryptonectria
; ORGANISM: Parasitica)
; STRAIN: EP713
; US-08-459-146-3

Query Match 23.6%; Score 47.5; DB 2; Length 3165;
Best Local Similarity 34.4%; Pred. No. 3.2e+02;
Matches 11; Conservative 8; Mismatches 8; Indels 5; Gaps 1;

QY 3 KWPFFAFHHQILATSTIAMSHCRPTVYKQTS 34
| | | | | : | : | : | | | | |
DB 1991 KYPAHFATFSELLAHNNVNV-----TEWKRLS 2017

RESULT 4
US-08-459-065-3
; Sequence 3, Application US/08459065
; Patent No. 5882642
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil Ho
; APPLICANT: Nuss, Donald Lee
; TITLE OF INVENTION: Genetically Engineered Transmissible
; TITLE OF INVENTION: Hypovirulence
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,065
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/832,117
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Catherine R
; REGISTRATION NUMBER: 34,240
; REFERENCE/DOCKET NUMBER: 8589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-6208
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3165 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Endothia parasitica (Cryptonectria
; ORGANISM: Parasitica)
; STRAIN: EP713
; US-08-459-065-3

Query Match 23.6%; Score 47.5; DB 2; Length 3165;
Best Local Similarity 34.4%; Pred. No. 3.2e+02;
Matches 11; Conservative 8; Mismatches 8; Indels 5; Gaps 1;

QY 3 KWPFFAFHHQILATSTIAMSHCRPTVYKQTS 34
| | | | | : | : | : | | | | |
DB 1991 KYPAHFATFSELLAHNNVNV-----TEWKRLS 2017

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2003, 17:24:23 ; Search time 41 Seconds
(without alignments)
117.001 Million cell updates/sec

Title: US-09-866-066-20

Perfect score: 201

Sequence: 1 PYKPFHFHAFHQILATSTAMSHCRPTVYKQTSII 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | ID | Description |
|------------|-------|-------------|------|--------------------------------|
| 1 | 201 | 100.0 | 36 | 23 AAE18025 Human ion channel, |
| 2 | 67 | 33.3 | 618 | 17 AAR97984 DmORF1 potassium c |
| 3 | 67 | 33.3 | 618 | 22 AA007616 Drosophila melanog |
| 4 | 67 | 33.3 | 1001 | 22 ABB58298 Drosophila melanog |
| 5 | 57 | 28.4 | 270 | 23 AAU91558 Amino acid sequenc |
| 6 | 55 | 27.4 | 97 | 23 AAU49085 Human ATP-depend |
| 7 | 55 | 27.4 | 97 | 23 AAU75354 Human Katal inhibi |
| 8 | 54.5 | 27.1 | 314 | 22 AAU24763 Human olfactory re |
| 9 | 54.5 | 27.1 | 314 | 23 ABP61148 Human GPCR protein |
| 10 | 54.5 | 27.1 | 314 | 23 AAU95760 Human olfactory an |

| | | | | |
|----|------|------|------|---------------------------------|
| 11 | 54.5 | 27.1 | 314 | 23 AAU85383 G-coupled olfactory |
| 12 | 52.5 | 26.1 | 223 | 21 AAY96675 Murine olfactory r |
| 13 | 52 | 25.9 | 128 | 22 AAU51853 Propionibacterium |
| 14 | 52 | 25.9 | 1373 | 22 ABG24466 Novel human diagno |
| 15 | 51 | 25.4 | 209 | 22 ABG09218 Novel human diagno |
| 16 | 50.5 | 25.1 | 32 | 22 ABG77020 Human colon cancer |
| 17 | 50.5 | 25.1 | 104 | 22 AAU28340 Novel human secret |
| 18 | 50 | 24.9 | 293 | 22 ABG05968 Novel human diagno |
| 19 | 50 | 24.9 | 512 | 23 ABB47839 Listeria monocytog |
| 20 | 49.5 | 24.6 | 102 | 22 AAB93368 Human protein sequ |
| 21 | 49 | 24.4 | 60 | 21 AAB51629 Human secreted pro |
| 22 | 49 | 24.4 | 66 | 22 AAM64236 Human brain expres |
| 23 | 49 | 24.4 | 68 | 22 AAU10542 Human polypeptide |
| 24 | 49 | 24.4 | 170 | 22 AAU29437 Human G protein-co |
| 25 | 49 | 24.4 | 170 | 23 ABG60725 Novel G protein co |
| 26 | 49 | 24.4 | 191 | 22 AAM24413 Human EST encoded |
| 27 | 49 | 24.4 | 241 | 21 AAB58199 Lung cancer associ |
| 28 | 48.5 | 24.1 | 251 | 22 AAG71611 Human olfactory re |
| 29 | 48.5 | 24.1 | 324 | 22 AAU24764 Human olfactory re |
| 30 | 48.5 | 24.1 | 324 | 23 AAU95759 Human olfactory an |
| 31 | 48.5 | 24.1 | 324 | 23 AAU95384 Olfactory G-couple |
| 32 | 48.5 | 24.1 | 485 | 22 AAG98251 Escherichia coli p |
| 33 | 48.5 | 24.1 | 1450 | 22 ABG10837 Novel human diagno |
| 34 | 48 | 23.9 | 74 | 21 AAB39002 Human secreted pro |
| 35 | 48 | 23.9 | 162 | 22 ABB28658 Peptide #1309 enco |
| 36 | 48 | 23.9 | 162 | 22 ABB33842 Peptide #1348 enco |
| 37 | 48 | 23.9 | 162 | 22 ABB19284 Protein #1283 enco |
| 38 | 48 | 23.9 | 162 | 22 AAM54608 Human brain expres |
| 39 | 48 | 23.9 | 162 | 22 AAM67014 Human bone marrow |
| 40 | 48 | 23.9 | 162 | 22 AAM14874 Peptide #1308 enco |
| 41 | 48 | 23.9 | 162 | 22 AAM27301 Peptide #1338 enco |
| 42 | 48 | 23.9 | 162 | 22 AAM2599 Peptide #1281 enco |
| 43 | 48 | 23.9 | 162 | 23 ABG36671 Human peptide enco |
| 44 | 48 | 23.9 | 425 | 22 ABG27006 Novel human diagno |
| 45 | 48 | 23.9 | 434 | 22 ABG12583 Novel human diagno |

ALIGNMENTS

RESULT 1

AAE18025
ID AAE18025 standard; peptide; 36 AA.

XX AAE18025;

AC AC
DT 07-MAY-2002 (first entry)

XX Human ion channel, Ion-157.

DE Human; ion channel; neurological disorder; psychiatric disorder;

KW schizophrenia; attention deficit hyperactivity disorder; depression;

KW proliferation disease; migraine; ischemia; neurodegenerative disease;

KW macular degeneration; Alzheimer's disease; congestive heart failure;

KW glaucoma; Parkinson's disease; cardiovascular disease; arrhythmia;

KW high blood pressure; restenosis; metabolic disease; neuroprotective;

KW obesity; hormonal disorder; polycystic ovarian syndrome; gene therapy;

XX alopecia; anxiety; stroke; neuroleptic; nootropic; cancer; diabetes.

OS Homo sapiens.

XX WO200192303-A2.

XX 06-DEC-2001.

XX 25-MAY-2001; 2001WO-US16967.

XX 26-MAY-2000; 2000US-207119P.

XX 26-MAY-2000; 2000US-207152P.

XX 26-MAY-2000; 2000US-207257P.

XX (PHAA) PHARMACIA & UPJOHN CO.

Query 1
Best Lo

Matches 11; Conservative 4; Mismatches 11; Indels 3; Gaps 1;

QY 1 PYKWPFFHFAFFHQILATSTIAMSHCRPTV 29
 II : I I I I I : : : :
 Db 153 PYSYSLMALFHTIL---TFRLSYCHSNI 178

RESULT 10

AAU95760
 ID AAU95760 standard; Protein; 314 AA.

AC AAU95760;

DT 02-JUL-2002 (first entry)

DE Human olfactory and pheromone G protein-coupled receptor #247.

DE Human; olfactory and pheromone G protein coupled; receptor;
 KW GPCR; tranquilizer; antidepressant; neuroleptic; endocrine; anabolic;
 KW anorectic; taste; fragrance; food additive; cosmetic; cell migration;
 KW sterility; psychotic disorder; neurological disorder; anxiety;
 KW schizophrenia; manic depression; depression; axonal growth;
 KW menstrual cycle; appetite sexual motivation; sexual attraction;
 KW aggression.

XX Homo sapiens.

XX WO200224726-A2.

PN 28-MAR-2002.

XX 21-SEP-2001; 2001WO-BE00162.

XX 22-SEP-2000; 2000EP-0870211.

XX (CHEM-) CHEMCOM SA.

PA Veithen A;

PI WPI: 2002-330013/36.

DR N-PSDB; ABR68647.

XX Novel pheromone G-protein coupled receptor and receptor-derived
 PT agonists, antagonists or inhibitors useful in food or cosmetic products
 PT or in the treatment or prevention of neurological disorders such as
 PT anxiety and schizophrenia

XX Disclosure; Page 753-754; 833pp; English.

XX The invention relates to olfactory and Pheromone G-protein coupled
 CC receptor (GPCR) or a protein 95% identical to the GPCR, a specific active
 CC portion and its encoding polynucleotide. Also included are an agonist,
 CC antagonist or inhibitor of the GPCR or the polynucleotide, a vector
 CC comprising the polynucleotide, a cell transformed by the vector, a non-
 CC human mammal comprising a partial or total deletion of the polynucleotide
 CC encoding the receptor and screening (detection and possibly, recovering)
 CC of compounds which are known or not known to be agonist, antagonists or
 CC inhibitors of natural compounds to the GPCR. The receptor-derived
 CC agonists, antagonists, inhibitors or compounds are used as an
 CC improvement, elimination or substitution of an existing taste and/or a
 CC fragrance of (or in) the food and/or cosmetic products. They can also be
 CC used in the preparation of medicament in the treatment and/or prevention
 CC of a mammalian disorder, such as cell migration, sterility, psychotic and
 CC neurological disorders, including anxiety, schizophrenia, manic
 CC depression, depression, for promoting axonal growth, nerve cell
 CC connection and nerve regeneration for modulating male and female
 CC endocrine functions, hormone production and the menstrual cycle, for the
 CC prevention or the treatment by stimulation of several mammalian
 CC behaviours, such as stimulation or suppression of appetite, sexual
 CC motivation, sexual attraction, aggression and for promoting or
 CC suppressing chemical communication between organisms. The present
 CC sequence is a human olfactory and pheromone GPCR protein sequence.

SQ Sequence 314 AA;

Query Match 27.1%; Score 54.5; DB 23; Length 314;
 Best Local Similarity 37.9%; Pred. No. 7.7;
 Matches 11; Conservative 4; Mismatches 11; Indels 3; Gaps 1;

QY 1 PYKWPFFHFAFFHQILATSTIAMSHCRPTV 29

II : I I I I I : : : :
 Db 153 PYSYSLMALFHTIL---TFRLSYCHSNI 178

RESULT 11

AAU85383
 ID AAU85383 standard; Protein; 314 AA.

XX AC AAU85383;

DT 08-MAY-2002 (first entry)

XX G-coupled olfactory receptor #244.

XX Human; olfactory G-coupled receptor; sensory perception of odourant;
 KW odour composition; taste composition.

XX Homo sapiens.

XX WO200198526-A2.

PN 27-DEC-2001.

XX 22-JUN-2001; 2001WO-US20122.

XX 22-JUN-2000; 2000US-213812P.

PR 13-MAR-2001; 2001US-0804291.

XX (SENO-) SENOMYX INC.

PA Zozulya S, Stryer L;

XX WPI: 2002-0833330/11.

DR N-PSDB; ABR37742.

XX Representing sensory perception of one or more odourants for the
 PT identification and design of tastes and odours comprises providing a
 PT representative group of n olfactory receptors

XX Claim 1; Page 168; 182pp; English.

XX The invention relates to a method of representing sensory perception of
 CC one or more odourants. The method comprises: (a) providing a
 CC representative class of n olfactory receptors or ligand binding domains
 CC (LBD) of these receptors; (b) measuring values X1 to Xn representative
 CC of at least one activity of one or more odourants selected from:
 CC (i) binding one or more odourants to the LBD of at least one of the n
 CC olfactory receptors; (ii) activating at least one of the n
 CC least one of the n olfactory receptors with the one or more odourants;
 CC and (c) generating a representation of sensory perception from the values
 CC X1 to Xn. The representation of the sensory perception of odourants is
 CC useful for the design and formulation of odour and taste compositions.
 CC AAU85140-AAU85393 represent human olfactory G-coupled receptor amino acid
 CC sequences of the invention.

SQ Sequence 314 AA;

Query Match 27.1%; Score 54.5; DB 23; Length 314;
 Best Local Similarity 37.9%; Pred. No. 7.7;
 Matches 11; Conservative 4; Mismatches 11; Indels 3; Gaps 1;

QY 1 PYKWPFFHFAFFHQILATSTIAMSHCRPTV 29

II : I I I I I : : : :
 Db 153 PYSYSLMALFHTIL---TFRLSYCHSNI 178

PS Example 45; Page 213; 250pp; English.

XX The present invention relates to the isolation of novel human genes
CC designated PHORI-A11 and PHORI-F5D6 and their encoded proteins. The
CC gene encoding PHORI-A11 maps to chromosome 1q23, and the gene encoding
CC PHORI-F5D6 maps to chromosome 7q33-q35. The PHORI-A11 and PHORI-F5D6
CC polynucleotide and polypeptide sequences are useful in diagnostic and
CC therapeutic methods, and compositions for various cancers such as
CC prostate cancer. The sequences are useful for inhibiting the growth of
CC cancer cells that express PHORI-A11 or PHORI-F5D6 and for treating
CC cancer. The PHORI-A11 or PHORI-F5D6 polypeptide or a fragment thereof
CC can be used to elicit an immune response. The present sequence
CC represents an amino acid sequence from the translation of the DNA
CC sequence for human PHORI-F5D6 splice variant B.

XX Sequence 270 AA;

Query Match 28.4%; Score 57; DB 23; Length 270;

Best Local Similarity 34.8%; Pred. No. 2.8; Mismatches 6; Indels 0; Gaps 0;

Matches 8; Conservative

OY 3 KWPFFHFAFFHQILATSTIAMSHC 25

DB 5 RWSFYHFWYDILSSHDVSNALC 27

RESULT 6

AM49085

ID AM49085 standard; Protein; 97 AA.

XX AC

XX AM49085;

XX O2-MAY-2002. (first entry)

XX Human ATP-dependent RNA helicase 11.

XX Human; ATP-dependent RNA helicase 11; recombinant production;

XX malignant tumour; cancer; blood disease; HIV infection; gene therapy;

XX human immunodeficiency virus; immune disorder; inflammatory condition;

XX cytosolic; anti-HIV; anti-inflammatory; immunomodulator; enzyme.

XX Homo sapiens.

XX WO200196534-A1.

XX 20-DEC-2001.

XX 14-MAY-2001; 2001WO-CN00744.

XX 19-MAY-2000; 2000CN-0115774.

XX (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.

XX Xie Y, Mao Y;

XX WPI; 2002-075589/10.

XX N-PSDB; ABA96858.

XX ATP-dependent human RNA helicase 11 and encoding polynucleotide, used

XX in diagnosis and treatment of malignant tumors, hemopathy, human

XX immunodeficiency virus infection, immunological diseases and

XX inflammation

XX Claim 1; Page 32; 40pp; Chinese.

XX This sequence represents human ATP-dependent RNA helicase 11. The

XX protein has a molecular weight of 11 kD. The invention relates to

XX human ATP-dependent RNA helicase 11 (AM49085), nucleic acids encoding it

XX (ABA96858), and a method for the recombinant production of ATP-dependent

XX RNA helicase 11. The present invention additionally discloses an

XX antagonist of ATP-dependent RNA helicase 11 for therapeutic use, and an

XX antibody which specifically binds to ATP-dependent RNA helicase 11.

XX ATP-dependent RNA helicase 11, and nucleotides which encode it may be

CC used for treating a variety of diseases, such as malignant tumours, blood
CC diseases, HIV (human immunodeficiency virus) infection, immune disorders
CC and inflammatory conditions. The protein may also be used to screen for
CC modulators of its activity or for peptide fingerprinting identification.
CC The polynucleotide can be used as a primer for nucleic acid amplification
CC reactions or as a probe for hybridisation reactions, or in producing gene
CC chips or microarrays.

SQ Sequence 97 AA;

Query Match 27.4%; Score 55; DB 23; Length 97;

Best Local Similarity 42.3%; Pred. No. 1.7;

Matches 11; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

OY 1 PYKPFHFAFFHQILATSTIAMSHCR 26

DB 28 PYKPMRFFIFSPILREKFSSESCR 53

RESULT 7

AAU75354

ID AAU75354 standard; Protein; 97 AA.

XX AC

XX AAU75354;

XX 23-APR-2002 (first entry)

XX Human Kazal inhibition factor 11.

XX Human; kazal inhibition factor 11; malignant tumour;

XX haemopathy; human immunodeficiency virus infection; HIV;

XX immunological disease; inflammation.

XX Homo sapiens.

XX WO200194539-A2.

XX 13-DEC-2001.

XX 28-MAY-2001; 2001WO-CN00872.

XX 26-MAY-2000; 2000CN-0115897.

XX (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.

XX Mao Y, Xie Y;

XX WPI; 2002-083228/11.

XX N-PSDB; ABA13294.

XX human Kazal inhibition factor 11 and encoding polynucleotide, used in

XX diagnosis and treatment of malignant tumours, haemopathy, human

XX immunodeficiency virus infection, immunological diseases and

XX inflammation

XX Claim 2; Page 30; 32pp; Chinese.

XX The invention relates to an isolated polypeptide of kazal inhibition

XX factor 11, the cDNA encoding it, and its fragment, analogue or

XX derivative. Also included are vectors expressing protein, a host cell

XX comprising the vector, the isolation of modulators of the protein and an

XX anti-kazal inhibition factor 11 antibody. The protein and nucleic acid

XX are used in diagnosis and treatment of a malignant tumour, haemopathy,

XX human immunodeficiency virus (HIV) infection, immunological diseases and

XX various inflammations. The present sequence represents the kazal

XX inhibition factor 11.

XX Sequence 97 AA;

Query Match 27.4%; Score 55; DB 23; Length 97;

Best Local Similarity 42.3%; Pred. No. 1.7;

Matches 11; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2003, 17:27:01 ; Search time 15 Seconds
(without alignments)
128.309 Million cell updates/sec

Title: US-09-866-066-20

Perfect score: 201

Sequence: 1 PKKWFHFAFHQILATSTIAMSHCRPTVYKQTSII 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/FCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/FCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 54.5 | 27.1 | 314 | 10 | US-09-886-055-495 |
| 2 | 49 | 24.4 | 66 | 10 | US-09-864-761-45919 |
| 3 | 49 | 24.4 | 170 | 10 | US-09-811-284-186 |
| 4 | 49 | 24.4 | 241 | 10 | US-09-925-302-537 |
| 5 | 48.5 | 24.1 | 324 | 10 | US-09-886-055-497 |
| 6 | 48.5 | 24.1 | 485 | 10 | US-09-741-669-299 |
| 7 | 48 | 23.9 | 162 | 10 | US-09-864-761-34582 |
| 8 | 47 | 23.4 | 197 | 10 | US-09-811-284-197 |
| 9 | 46 | 22.9 | 258 | 10 | US-09-746-491-61 |
| 10 | 46 | 22.9 | 394 | 10 | US-09-939-484-4 |
| 11 | 46 | 22.9 | 394 | 10 | US-09-939-483-4 |
| 12 | 46 | 22.9 | 394 | 10 | US-09-798-584-15 |
| 13 | 46 | 22.9 | 405 | 10 | US-09-939-484-5 |
| 14 | 46 | 22.9 | 405 | 10 | US-09-939-483-5 |
| 15 | 45.5 | 22.6 | 291 | 10 | US-09-815-242-5332 |
| 16 | 45.5 | 22.6 | 366 | 10 | US-09-815-242-12345 |
| 17 | 45.5 | 22.6 | 366 | 10 | US-09-815-242-12345 |
| 18 | 45 | 22.4 | 107 | 9 | US-09-764-904-62 |
| 19 | 45 | 22.4 | 107 | 9 | US-10-091-548-62 |

| | | | | | | |
|----|------|------|------|----|---------------------|--------------------|
| 20 | 45 | 22.4 | 107 | 10 | US-09-764-860-580 | Sequence 580, App |
| 21 | 45 | 22.4 | 169 | 9 | US-09-738-626-5549 | Sequence 5549, Ap |
| 22 | 45 | 22.4 | 330 | 10 | US-09-746-491-59 | Sequence 59, Appl |
| 23 | 45 | 22.4 | 365 | 10 | US-09-746-491-60 | Sequence 60, Appl |
| 24 | 45 | 22.4 | 374 | 10 | US-09-798-584-1 | Sequence 1, Appl |
| 25 | 45 | 22.4 | 791 | 9 | US-09-738-626-4530 | Sequence 4530, Ap |
| 26 | 45 | 22.4 | 1091 | 9 | US-09-423-126-2 | Sequence 2, Appl |
| 27 | 45 | 22.4 | 1091 | 10 | US-09-280-197-2 | Sequence 2, Appl |
| 28 | 44.5 | 22.1 | 63 | 10 | US-09-864-761-44430 | Sequence 44430, A |
| 29 | 44 | 21.9 | 330 | 10 | US-09-746-491-58 | Sequence 58, Appl |
| 30 | 44 | 21.9 | 332 | 9 | US-10-001-835-140 | Sequence 140, Appl |
| 31 | 44 | 21.9 | 339 | 12 | US-10-067-291-1 | Sequence 1, Appl |
| 32 | 44 | 21.9 | 355 | 9 | US-09-922-895-1 | Sequence 1, Appl |
| 33 | 44 | 21.9 | 355 | 10 | US-09-931-381A-16 | Sequence 16, Appl |
| 34 | 44 | 21.9 | 355 | 12 | US-10-106-823-4 | Sequence 4, Appl |
| 35 | 44 | 21.9 | 393 | 10 | US-09-746-491-22 | Sequence 22, Appl |
| 36 | 44 | 21.9 | 1333 | 10 | US-09-815-242-10936 | Sequence 10936, A |
| 37 | 44 | 21.9 | 2310 | 10 | US-09-995-542-10 | Sequence 10, Appl |
| 38 | 43.5 | 21.6 | 233 | 9 | US-09-974-879-398 | Sequence 398, App |
| 39 | 43.5 | 21.6 | 791 | 9 | US-09-967-386-1 | Sequence 1, Appl |
| 40 | 43.5 | 21.6 | 810 | 10 | US-09-946-893-2 | Sequence 2, Appl |
| 41 | 43.5 | 21.6 | 1169 | 9 | US-09-870-759-126 | Sequence 126, App |
| 42 | 43 | 21.4 | 355 | 10 | US-09-789-482-4 | Sequence 4, Appl |
| 43 | 43 | 21.4 | 355 | 10 | US-09-789-486-4 | Sequence 4, Appl |
| 44 | 43 | 21.4 | 360 | 9 | US-10-120-394-20 | Sequence 20, Appl |
| 45 | 43 | 21.4 | 360 | 9 | US-09-764-413-20 | Sequence 20, Appl |

ALIGNMENTS

RESULT 1
US-09-886-055-495
; Sequence 495, Application US/09886055
; Patent No. US2002013273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 495
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-055-495

Query Match 27.1%; Score 54.5; DB 10; Length 314;
Best Local Similarity 37.9%; Pred. No. 1.9;
Matches 11; Conservative 4; Mismatches 11; Indels 3; Gaps 1;
OY 1 PKKWFHFAFHQILATSTIAMSHCRPTV 29
Db 153 PYSYSFLMALPHTIL---TFRLSYCHSNI 178

RESULT 2
US-09-864-761-45919
; Sequence 45919, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

US-09-886-055-497
 ; Sequence 497, Application US/09886055
 ; Patent No. US20020132273A1
 ; GENERAL INFORMATION:
 ; APPLICANT: STRYER, LOBERT
 ; APPLICANT: ZOZULYA, SERGEY
 ; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
 ; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
 ; FILE REFERENCE: 078003-0277150
 ; CURRENT APPLICATION NUMBER: US/09/886,055
 ; CURRENT FILING DATE: 2001-06-22
 ; PRIOR APPLICATION NUMBER: 60/213,812
 ; PRIOR FILING DATE: 2000-06-22
 ; NUMBER OF SEQ ID NOS: 522
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 497
 ; LENGTH: 324
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-886-055-497

Query Match 24.1%; Score 48.5; DB 10; Length 324;
 Best Local Similarity 31.0%; Pred. No. 16;
 Matches 9; Conservative 7; Mismatches 10; Indels 3; Gaps 1;

QY 1 PYKWPFFHFAFFHOILATSIAMSHCRPTV 29
 || : | | | : | : : | | :
 DB 148 PPIYSLVALFHTVI---TFRITYCGPNL 173

RESULT 6
 US-09-741-669-299
 ; Sequence 299, Application US/09741669
 ; Patent No. US20020022719A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Forsyth, R. Allyn
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; TITLE OF INVENTION: Genes identified as required for
 ; TITLE OF INVENTION: Proliferation of E. coli
 ; FILE REFERENCE: ELITRA-009A
 ; CURRENT APPLICATION NUMBER: US/09/741,669
 ; CURRENT FILING DATE: 2000-12-19
 ; PRIOR APPLICATION NUMBER: US 60/173005
 ; PRIOR FILING DATE: 1999-12-23
 ; NUMBER OF SEQ ID NOS: 481
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 299
 ; LENGTH: 485
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-09-741-669-299

Query Match 24.1%; Score 48.5; DB 10; Length 485;
 Best Local Similarity 50.0%; Pred. No. 25;
 Matches 10; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 1 PYKWPFFHFAFFHOILATSIYI 20
 | : | | : | | | | |
 DB 266 PFYGFSEFS-FHQFLVVGTI 284

RESULT 7
 US-09-864-761-34582
 ; Sequence 34582, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aecomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263,6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/006666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/006667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/006668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/006669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/006665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/006668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/006663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/006662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/006661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/006670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 34582
 ; LENGTH: 162
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC000029.14
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.1
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.7
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.3
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.4
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.5
 ; OTHER INFORMATION: SWISSPROT HIT: P09053, EVALUATE 4.90e+00
 ; OTHER INFORMATION: EST_HUMAN HIT: BF104006.1, EVALUATE 4.00e-01
 US-09-864-761-34582

Query Match 23.9%; Score 48; DB 10; Length 162;
 Best Local Similarity 37.5%; Pred. No. 8.6;
 Matches 12; Conservative 4; Mismatches 12; Indels 4; Gaps 2;

QY 1 PYKWPFFHFAFFHOILATSIAMSHCRPT 28
 | | | : | : | | : | | |
 DB 115 PMGWPSYPLCFPMWHLWELTSAFSLSHARPT 146

RESULT 8
 US-09-811-284-197
 ; Sequence 197, Application US/09811284

2007

; FILE REFERENCE: ELITRA.01

```
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5332
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5332
```

```
Query Match      22.6%; Score 45.5; DB 10; Length 291;
Best Local Similarity 36.7%; Pred. No. 39;
Matches 11; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

QY      2 YKWPFFHFAFPHOILATSTIAMSHCRPTVYK 31
      ||| | | | | | | | | | | | | | |
Db     122 YKWFAYPDEH-IETVNLTIKDPVYFK 150
```

```
Search completed: March 24, 2003, 17:27:40
Job time : 16 secs
```

```

; Patent No. US20020058306A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; TITLE OF INVENTION: No. US20020058306A1el G Protein-Coupled Receptors
; FILE REFERENCE: 00167US1
; CURRENT FILING DATE: 2001-03-16
; CURRENT APPLICATION NUMBER: US/09/811,284
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,783
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,907
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,918
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,960
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,917
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/192,945
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,916
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,923
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,933
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,830
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,234
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,155
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,935
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 197
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-811-284-197

Query Match      23.4%; Score 47; DB 10; Length 197;
Best Local Similarity 37.5%; Pred. No. 15;
Matches 9; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY      4 WPFHFAFFHQLATSTIAMSHCRP 27
DB     135 WPELFSSFFSLSKELILAEFWSP 158

; Patent No. US20020058306A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20020137202A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-621
; CURRENT APPLICATION NUMBER: US/09/746,491
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: USSN 60/171,329
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 61
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-746-491-61

Query Match      22.9%; Score 46; DB 10; Length 258;
Best Local Similarity 31.2%; Pred. No. 29;
Matches 10; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

```

```

QY      1 PYK----WPFHFAFFHQLATSTIAMSHCRPT 28
DB     71 PHKAGVQWRFGSFYFAITVITTTIGYGHAAPS 102

; Patent No. US20020032322A1
; GENERAL INFORMATION:
; APPLICANT: Duprat, Fabrice
; APPLICANT: Lesage, Florian
; APPLICANT: Fink, Michel
; APPLICANT: Lazdunski, Michel
; TITLE OF INVENTION: FAMILY OF MAMMALIAN POTASSIUM CHANNELS, THEIR CLONING
; FILE REFERENCE: 1201-CIP-DIV-00
; CURRENT APPLICATION NUMBER: US/09/939,484
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 09/144,914
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 08/749,816
; PRIOR FILING DATE: 1996-11-15
; PRIOR APPLICATION NUMBER: 60/095,234
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: FR 96/01565
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: TASK
; US-09-939-484-4

Query Match      22.9%; Score 46; DB 10; Length 394;
Best Local Similarity 31.2%; Pred. No. 46;
Matches 10; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

QY      1 PYK----WPFHFAFFHQLATSTIAMSHCRPT 28
DB     71 PHKAGVQWRFGSFYFAITVITTTIGYGHAAPS 102

; Patent No. US20020032322A1
; GENERAL INFORMATION:
; APPLICANT: Duprat, Fabrice
; APPLICANT: Lesage, Florian
; APPLICANT: Fink, Michel
; APPLICANT: Lazdunski, Michel
; TITLE OF INVENTION: FAMILY OF MAMMALIAN POTASSIUM CHANNELS, THEIR CLONING
; FILE REFERENCE: 1201-CIP-DIV-2-00
; CURRENT APPLICATION NUMBER: US/09/939,483
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 09/144,914
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 08/749,816
; PRIOR FILING DATE: 1996-11-15
; PRIOR APPLICATION NUMBER: 60/095,234
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: FR 96/01565
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 394

```

```

; FILE REFERENCE: Acomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45919
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC026206.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 0.51
; OTHER INFORMATION: EST_HUMAN HIT: BF340649.1, EVALUE 7.60e-02
; OTHER INFORMATION: SWISSPROT HIT: P13597, EVALUE 3.40e+00
US-09-864-761-45919

```

```

Query Match      24.4%; Score 49; DB 10; Length 66;
Best Local Similarity 48.0%; Pred. No. 2.2;
Matches 12; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

```

```

QY 12 HQILATSTIAMSHCRPTVYKQTSII 36
| | | | | | | | | | | | | | | |
DB 36 HQGLSTEMKASHCLPLKSKRWCVII 60
| | | | | | | | | | | | | | | |

```

```

RESULT 3
US-09-811-284-186
; Sequence 186, Application US/09811284
; Patent No. US20020058306A1
; GENERAL INFORMATION:
; APPLICANT: Vogel, Gabriel
; TITLE OF INVENTION: No. US20020058306A1e1 G Protein-Coupled Receptors
; FILE REFERENCE: 00167US1
; CURRENT APPLICATION NUMBER: US/09/811,284
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,783
; PRIOR FILING DATE: 2000-03-16

```

```

; PRIOR APPLICATION NUMBER: 60/189,907
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,918
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,960
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,917
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/192,945
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,916
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,923
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,933
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,830
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,234
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,155
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,935
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 186
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-284-186

```

```

Query Match      24.4%; Score 49; DB 10; Length 170;
Best Local Similarity 28.6%; Pred. No. 6.4;
Matches 12; Conservative 11; Mismatches 9; Indels 10; Gaps 2;

```

```

QY 3 KWPFFAFHQILATS-----TIAMSHCRPTVYKQTSII 36
| | | | | | | | | | | | | | | |
DB 28 KWPFQSSW--LSSTSPCHTCIOITLSVHCRNLINKSLVI 67
| | | | | | | | | | | | | | | |

```

```

RESULT 4
US-09-925-302-537
; Sequence 537, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 537
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-537

```

```

Query Match      24.4%; Score 49; DB 10; Length 241;
Best Local Similarity 35.5%; Pred. No. 9.5;
Matches 11; Conservative 5; Mismatches 13; Indels 2; Gaps 1;

```

```

QY 4 WPFHAFHQILATSTIAMSHCRPTVYKQTS 34
| | | | | | | | | | | | | | | |
DB 19 WPFYLEY--SLLAEFTLVVKOKLPGVYVOPS 47
| | | | | | | | | | | | | | | |

```

```

RESULT 5

```